

CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number  
WO 02/046415 A2(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
C12Q 1/68, C07K 14/705, 16/28, A61K 31/17, G01N  
33/50, 33/68

(21) International Application Number: PCT/US01/46963

(22) International Filing Date: 5 December 2001 (05.12.2001)

(25) Filing Language: English

(26) Publication Language: English

## (30) Priority Data:

60/254,303	8 December 2000 (08.12.2000)	US
60/256,190	15 December 2000 (15.12.2000)	US
60/257,504	21 December 2000 (21.12.2000)	US
60/261,546	12 January 2001 (12.01.2001)	US
60/262,832	19 January 2001 (19.01.2001)	US
60/264,377	26 January 2001 (26.01.2001)	US
60/266,019	2 February 2001 (02.02.2001)	US

(71) Applicant (for all designated States except US): INCYTE  
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94304 (US).

## (72) Inventors; and

(75) Inventors/Applicants (for US only): LEE, Ernestine,  
A. [US/US]; 624 Kains Street, Albany, CA 94706 (US).  
BAUGHN, Maria, R. [US/US]; 14244 Santiago Road, San  
Leandro, CA 94577 (US). YUE, Henry [US/US]; 826 Lois  
Avenue, Sunnyvale, CA 94087 (US). DING, Li [CN/US];  
3353 Alma Street #146, Palo Alto, CA 94306 (US). RAU-  
MANN, Brigitte, E. [US/US]; 5801 South Dorchester  
Avenue #3B, Chicago, IL 60637 (US). HAFALIA, April,  
J., A. [US/US]; 2227 Calle de Primavera, Santa Clara, CA  
95054 (US). KHAN, Farrah, A. [IN/US]; 9445 Harrison  
Street, Des Plaines, IL 60016 (US). NGUYEN, Dannie,  
B. [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118  
(US). ELLIOTT, Vicki, S. [US/US]; 3770 Polten Place  
Way, San Jose, CA 95121 (US). RAMKUMAR, Jay-  
alaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA  
94555 (US). WALLA, Narinder, K. [US/US]; 890 Davis  
Street #205, San Leandro, CA 94577 (US). ISON, Craig,  
H. [US/US]; 1242 Weathersfield Way, San Jose, CA 95118  
(US). LU, Yan [CN/US]; 3885 Corrina Way, Palo Alto,  
CA 94303 (US). GANDHI, Ameena, R. [US/US]; 705 5th  
Avenue, San Francisco, CA 94118 (US). WARREN, Brid-  
get, A. [US/US]; 10130 Parkwood Drive #2, Cupertino,CA 95014 (US). DUGGAN, Brendan, M. [AU/US]; 243  
Buena Vista Avenue #306, Sunnyvale, CA 94086 (US).  
TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee  
Street #5, San Francisco, CA 94107 (US). BURFORD,  
Neil [GB/US]; 105 Wildwood Circle, Durham, CT 06422  
(US). LU, Dyung, Alna, M. [US/US]; 233 Coy Drive,  
San Jose, CA 95123 (US). LAL, Preeti, G. [IN/US]; P.O.  
Box 5142, Santa Clara, CA 95056 (US). YAO, Monique,  
G. [US/US]; 1189 Woodgate Drive, Carmel, IN 46033  
(US). XU, Yuming [US/US]; 1739 Walnut Drive, Moun-  
tain View, CA 94040 (US). BRUNS, Christopher, M.  
[US/US]; 575 S. Rengstorff Avenue #126, Mountain View,  
CA 94040 (US). THANGAVELU, Kavitha [IN/US];  
1950 Montecito Avenue #23, Mountain View, CA 94043  
(US). SWARNAKAR, Anita [CA/US]; 8 Locksley Ave-  
nue #5D, San Francisco, CA 94122 (US). TANG, Y, Tom  
[US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).  
AZIMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive,  
Castro Valley, CA 94552 (US). THORNTON, Michael  
[US/US]; 9 Medway Road, Woodside, CA 94062-2612  
(US). ARVIZU, Chandra [US/US]; 490 Sherwood Way  
#1, Menlo Park, CA 94025 (US). POLICKY, Jennifer, L.  
[US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,  
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,  
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,  
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).

## Published:

— without international search report and to be republished  
upon receipt of that report

[Continued on next page]

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: The invention provides human transporters and ion channels (TRICH) and polynucleotides which identify and en-  
code TRICH. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also  
provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of TRICH.

WO 02/046415 A2



**(48) Date of publication of this corrected version:**

6 September 2002

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**(15) Information about Correction:**

see PCT Gazette No. 36/2002 of 6 September 2002, Section II

## TRANSPORTERS AND ION CHANNELS

This invention relates to nucleic acid and amino acid sequences of transporters and ion channels and to the use of these sequences in the diagnosis, treatment, and prevention of transport, neurological, muscle, immunological and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

## BACKGROUND OF THE INVENTION

Eukaryotic cells are surrounded and subdivided into functionally distinct organelles by hydrophobic lipid bilayer membranes which are highly impermeable to most polar molecules. Cells and organelles require transport proteins to import and export essential nutrients and metal ions including  $K^+$ ,  $NH_4^+$ ,  $P_i$ ,  $SO_4^{2-}$ , sugars, and vitamins, as well as various metabolic waste products. Transport proteins also play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). Transport can occur by a passive concentration-dependent mechanism, or can be linked to an energy source such as ATP hydrolysis or an ion gradient. Proteins that function in transport include carrier proteins, which bind to a specific solute and undergo a conformational change that translocates the bound solute across the membrane, and channel proteins, which form hydrophilic pores that allow specific solutes to diffuse through the membrane down an electrochemical solute gradient.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous  $Na^+/K^+$  ATPase system. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging techniques and for specific targeting of radioisotopes to the thyroid gland (Levy,

O. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) *J. Biol. Chem.* 273:7501-7506).

One of the largest families of transporters is the major facilitator superfamily (MFS), also  
5 called the uniporter-symporter-antiporter family. MFS transporters are single polypeptide carriers that transport small solutes in response to ion gradients. Members of the MFS are found in all classes of living organisms, and include transporters for sugars, oligosaccharides, phosphates, nitrates, nucleosides, monocarboxylates, and drugs. MFS transporters found in eukaryotes all have a structure comprising 12 transmembrane segments (Pao, S.S. et al. (1998) *Microbiol. Molec. Biol.*  
10 *Rev.* 62:1-34). The largest family of MFS transporters is the sugar transporter family, which includes the seven glucose transporters (GLUT1-GLUT7) found in humans that are required for the transport of glucose and other hexose sugars. These glucose transport proteins have unique tissue distributions and physiological functions. GLUT1 provides many cell types with their basal glucose requirements and transports glucose across epithelial and endothelial barrier tissues; GLUT2  
15 facilitates glucose uptake or efflux from the liver; GLUT3 regulates glucose supply to neurons; GLUT4 is responsible for insulin-regulated glucose disposal; and GLUT5 regulates fructose uptake into skeletal muscle. Defects in glucose transporters are involved in a recently identified neurological syndrome causing infantile seizures and developmental delay, as well as glycogen storage disease, Fanconi-Bickel syndrome, and non-insulin-dependent diabetes mellitus (Mueckler,  
20 M. (1994) *Eur. J. Biochem.* 219:713-725; Longo, N. and L.J. Elsas (1998) *Adv. Pediatr.* 45:293-313).

Monocarboxylate anion transporters are proton-coupled symporters with a broad substrate specificity that includes L-lactate, pyruvate, and the ketone bodies acetate, acetoacetate, and beta-hydroxybutyrate. At least seven isoforms have been identified to date. The isoforms are  
25 predicted to have twelve transmembrane (TM) helical domains with a large intracellular loop between TM6 and TM7, and play a critical role in maintaining intracellular pH by removing the protons that are produced stoichiometrically with lactate during glycolysis. The best characterized H<sup>+</sup>-monocarboxylate transporter is that of the erythrocyte membrane, which transports L-lactate and a wide range of other aliphatic monocarboxylates. Other cells possess H<sup>+</sup>-linked monocarboxylate  
30 transporters with differing substrate and inhibitor selectivities. In particular, cardiac muscle and tumor cells have transporters that differ in their K<sub>m</sub> values for certain substrates, including stereoselectivity for L- over D-lactate, and in their sensitivity to inhibitors. There are Na<sup>+</sup>-monocarboxylate cotransporters on the luminal surface of intestinal and kidney epithelia, which allow the uptake of lactate, pyruvate, and ketone bodies in these tissues. In addition, there are  
35 specific and selective transporters for organic cations and organic anions in organs including the



kidney, intestine and liver. Organic anion transporters are selective for hydrophobic, charged molecules with electron-attracting side groups. Organic cation transporters, such as the ammonium transporter, mediate the secretion of a variety of drugs and endogenous metabolites, and contribute to the maintenance of intercellular pH (Poole, R.C. and A.P. Halestrap (1993) *Am. J. Physiol.*

- 5 264:C761-C782; Price, N.T. et al. (1998) *Biochem. J.* 329:321-328; and Martinelle, K. and I. Haggstrom (1993) *J. Biotechnol.* 30:339-350).

ATP-binding cassette (ABC) transporters are members of a superfamily of membrane proteins that transport substances ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC  
 10 transporters consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-  
 15 molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic  
 20 hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) *Meth. Enzymol.* 292:130-162).

A number of metal ions such as iron, zinc, copper, cobalt, manganese, molybdenum,  
 25 selenium, nickel, and chromium are important as cofactors for a number of enzymes. For example, copper is involved in hemoglobin synthesis, connective tissue metabolism, and bone development, by acting as a cofactor in oxidoreductases such as superoxide dismutase, ferroxidase (ceruloplasmin), and lysyl oxidase. Copper and other metal ions must be provided in the diet, and are absorbed by transporters in the gastrointestinal tract. Plasma proteins transport the metal ions to  
 30 the liver and other target organs, where specific transporters move the ions into cells and cellular organelles as needed. Imbalances in metal ion metabolism have been associated with a number of disease states (Danks, D.M. (1986) *J. Med. Genet.* 23:99-106).

P-type ATPases comprise a class of cation-transporting transmembrane proteins. They are integral membrane proteins which use an aspartyl phosphate intermediate to move cations across a  
 35 membrane. Features of P-type ATPases include: (i) a cation channel; (ii) a stalk, formed by

extensions of the transmembrane  $\alpha$ -helices into the cytoplasm; (iii) an ATP binding domain; (iv) a phosphorylated aspartic acid; (v) an adjacent transduction domain; (vi) a phosphatase domain, which removes the phosphate from the aspartic acid as part of the reaction cycle; and (vii) six or more transmembrane domains. Included in this class are heavy metal-transporting ATPases as well  
5 as aminophospholipid transporters.

The transport of phosphatidylserine and phosphatidylethanolamine by aminophospholipid translocase results in the movement of these molecules from one side of a bilayer to another. This transport is conducted by a newly identified subfamily of P-type ATPases which are proposed to be amphipath transporters. Amphipath transporters move molecules having both a hydrophilic and a  
10 hydrophobic region. As many as seventeen different genes belong to this P-type ATPases subfamily, being grouped into several distinct classes and subclasses (Halleck, M.S. et al., (1999) *Physiol. Genomics* 1:139-150; Vulpe, C. et al., (1993) *Nat. Genet.* 3:7-13).

Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant  
15 fraction of fatty acid transport appears to occur via a high affinity, low capacity protein-mediated transport process. Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates uptake of long-chain fatty  
20 acids (Hui, T.Y. et al. (1998) *J. Biol. Chem.* 273:27420-27429).

The lipocalin superfamily constitutes a phylogenetically conserved group of more than forty proteins that function as extracellular ligand-binding proteins which bind and transport small hydrophobic molecules. Members of this family function as carriers of retinoids, odorants, chromophores, pheromones, allergens, and sterols, and in a variety of processes including nutrient  
25 transport, cell growth regulation, immune response, and prostaglandin synthesis. A subset of these proteins may be multifunctional, serving as either a biosynthetic enzyme or as a specific enzyme inhibitor. (Tanaka, T. et al. (1997) *J. Biol. Chem.* 272:15789-15795; and van't Hof, W. et al. (1997) *J. Biol. Chem.* 272:1837-1841.)

Members of the lipocalin family display unusually low levels of overall sequence  
30 conservation. Pairwise sequence identity often falls below 20%. Sequence similarity between family members is limited to conserved cysteines which form disulfide bonds and three motifs which form a juxtaposed cluster that functions as a target cell recognition site. The lipocalins share an eight stranded, anti-parallel beta-sheet which folds back on itself to form a continuously hydrogen-bonded beta-barrel. The pocket formed by the barrel functions as an internal ligand  
35 binding site. Seven loops (L1 to L7) form short beta-hairpins, except loop L1 which is a large

omega loop that forms a lid to partially close the internal ligand-binding site (Flower (1996) Biochem. J. 318:1-14).

Lipocalins are important transport molecules. Each lipocalin associates with a particular ligand and delivers that ligand to appropriate target sites within the organism. Retinol-binding protein (RBP), one of the best characterized lipocalins, transports retinol from stores within the liver to target tissues. Apolipoprotein D (apo D), a component of high density lipoproteins (HDLs) and low density lipoproteins (LDLs), functions in the targeted collection and delivery of cholesterol throughout the body. Lipocalins are also involved in cell regulatory processes. Apo D, which is identical to gross-cystic-disease-fluid protein (GCDFP)-24, is a progesterone/pregnenolone-binding protein expressed at high levels in breast cyst fluid. Secretion of apo D in certain human breast cancer cell lines is accompanied by reduced cell proliferation and progression of cells to a more differentiated phenotype. Similarly, apo D and another lipocalin,  $\alpha_1$ -acid glycoprotein (AGP), are involved in nerve cell regeneration. AGP is also involved in anti-inflammatory and immunosuppressive activities. AGP is one of the positive acute-phase proteins (APP); circulating levels of AGP increase in response to stress and inflammatory stimulation. AGP accumulates at sites of inflammation where it inhibits platelet and neutrophil activation and inhibits phagocytosis. The immunomodulatory properties of AGP are due to glycosylation. AGP is 40% carbohydrate, making it unusually acidic and soluble. The glycosylation pattern of AGP changes during acute-phase response, and deglycosylated AGP has no immunosuppressive activity (Flower (1994) FEBS Lett. 354:7-11; Flower (1996) *supra*).

The lipocalin superfamily also includes several animal allergens, including the mouse major urinary protein (mMUP), the rat  $\alpha$ -2-microglobulin (rA2U), the bovine  $\beta$ -lactoglobulin ( $\beta$ lg), the cockroach allergen (Bla g4), bovine dander allergen (Bos d2), and the major horse allergen, designated *Equus caballus* allergen 1 (Equ c1). Equ c1 is a powerful allergen responsible for about 80% of anti-horse IgE antibody response in patients who are chronically exposed to horse allergens. It appears that lipocalins may contain a common structure that is able to induce the IgE response (Gregoire, C. et al., (1996) J. Biol. Chem. 271:32951-32959).

Lipocalins are used as diagnostic and prognostic markers in a variety of disease states. The plasma level of AGP is monitored during pregnancy and in diagnosis and prognosis of conditions including cancer chemotherapy, renal disfunction, myocardial infarction, arthritis, and multiple sclerosis. RBP is used clinically as a marker of tubular reabsorption in the kidney, and apo D is a marker in gross cystic breast disease (Flower (1996) *supra*). Additionally, the use of lipocalin animal allergens may help in the diagnosis of allergic reactions to horses (Gregoire *supra*), pigs, cockroaches, mice and rats.

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions

and charged metabolites between the cytosol and the mitochondrial matrix. Examples include the ADP, ATP carrier protein; the 2-oxoglutarate/malate carrier; the phosphate carrier protein; the pyruvate carrier; the dicarboxylate carrier which transports malate, succinate, fumarate, and phosphate; the tricarboxylate carrier which transports citrate and malate; and the Grave's disease carrier protein, a protein recognized by IgG in patients with active Grave's disease, an autoimmune disorder resulting in hyperthyroidism. Proteins in this family consist of three tandem repeats of an approximately 100 amino acid domain, each of which contains two transmembrane regions (Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, p. 551; PROSITE PDOC00189 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) \*275000 Graves Disease).

This class of transporters also includes the mitochondrial uncoupling proteins, which create proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. The result is energy dissipation in the form of heat. Mitochondrial uncoupling proteins have been implicated as modulators of thermoregulation and metabolic rate, and have been proposed as potential targets for drugs against metabolic diseases such as obesity (Ricquier, D. et al. (1999) *J. Int. Med.* 245:637-642).

#### **Ion Channels**

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form ion-selective pores within the membrane. There are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

#### **Ion Transporters**

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient. These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including Na<sup>+</sup>-K<sup>+</sup> ATPase, Ca<sup>2+</sup>-ATPase, and H<sup>+</sup>-ATPase, are activated by a phosphorylation event. P-class ion transporters, also known as E1-E2 type ATPases, are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na<sup>+</sup> and Ca<sup>2+</sup> are low and cytosolic concentration of K<sup>+</sup> is high. The vacuolar (V) class of ion transporters includes H<sup>+</sup> pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are

responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of  $H^+$  pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate ( $P_i$ ).

The P-ATPases are hexamers of a 100 kD subunit with ten transmembrane domains and several large cytoplasmic regions that may play a role in ion binding (Scarborough, G.A. (1999) Curr. Opin. Cell Biol. 11:517-522). The V-ATPases are composed of two functional domains: the  $V_1$  domain, a peripheral complex responsible for ATP hydrolysis; and the  $V_0$  domain, an integral complex responsible for proton translocation across the membrane. The F-ATPases are structurally and evolutionarily related to the V-ATPases. The F-ATPase  $F_0$  domain contains 12 copies of the c subunit, a highly hydrophobic protein composed of two transmembrane domains and containing a single buried carboxyl group in TM2 that is essential for proton transport. The V-ATPase  $V_0$  domain contains three types of homologous c subunits with four or five transmembrane domains and the essential carboxyl group in TM4 or TM3. Both types of complex also contain a single a subunit that may be involved in regulating the pH dependence of activity (Forgacs, M. (1999) J. Biol. Chem. 274:12951-12954).

The resting potential of the cell is utilized in many processes involving carrier proteins and gated ion channels. Carrier proteins utilize the resting potential to transport molecules into and out of the cell. Amino acid and glucose transport into many cells is linked to sodium ion co-transport (symport) so that the movement of  $Na^+$  down an electrochemical gradient drives transport of the other molecule up a concentration gradient. Similarly, cardiac muscle links transfer of  $Ca^{2+}$  out of the cell with transport of  $Na^+$  into the cell (antiport).

#### Gated Ion Channels

Gated ion channels control ion flow by regulating the opening and closing of pores. The ability to control ion flux through various gating mechanisms allows ion channels to mediate such diverse signaling and homeostatic functions as neuronal and endocrine signaling, muscle contraction, fertilization, and regulation of ion and pH balance. Gated ion channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open their pores in response to mechanical stress; voltage-gated channels (e.g.,  $Na^+$ ,  $K^+$ ,  $Ca^{2+}$ , and  $Cl^-$  channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g., acetylcholine-, serotonin-, and glutamate-gated cation channels, and GABA- and glycine-gated chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post translational modifications, such as phosphorylation.

Mechanically-gated or mechanosensitive ion channels act as transducers for the senses of

touch, hearing, and balance, and also play important roles in cell volume regulation, smooth muscle contraction, and cardiac rhythm generation. A stretch-inactivated channel (SIC) was recently cloned from rat kidney. The SIC channel belongs to a group of channels which are activated by pressure or stress on the cell membrane and conduct both  $\text{Ca}^{2+}$  and  $\text{Na}^+$  (Suzuki, M. et al. (1999) J.

5 Biol. Chem. 274:6330-6335).

The pore-forming subunits of the voltage-gated cation channels form a superfamily of ion channel proteins. The characteristic domain of these channel proteins comprises six transmembrane domains (S1-S6), a pore-forming region (P) located between S5 and S6, and intracellular amino and carboxy termini. In the  $\text{Na}^+$  and  $\text{Ca}^{2+}$  subfamilies, this domain is repeated four times, while in the

10  $\text{K}^+$  channel subfamily, each channel is formed from a tetramer of either identical or dissimilar subunits. The P region contains information specifying the ion selectivity for the channel. In the case of  $\text{K}^+$  channels, a GYG tripeptide is involved in this selectivity (Ishii, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656).

Voltage-gated  $\text{Na}^+$  and  $\text{K}^+$  channels are necessary for the function of electrically excitable

15 cells, such as nerve and muscle cells. Action potentials, which lead to neurotransmitter release and muscle contraction, arise from large, transient changes in the permeability of the membrane to  $\text{Na}^+$  and  $\text{K}^+$  ions. Depolarization of the membrane beyond the threshold level opens voltage-gated  $\text{Na}^+$  channels. Sodium ions flow into the cell, further depolarizing the membrane and opening more voltage-gated  $\text{Na}^+$  channels, which propagates the depolarization down the length of the cell.

20 Depolarization also opens voltage-gated potassium channels. Consequently, potassium ions flow outward, which leads to repolarization of the membrane. Voltage-gated channels utilize charged residues in the fourth transmembrane segment (S4) to sense voltage change. The open state lasts only about 1 millisecond, at which time the channel spontaneously converts into an inactive state that cannot be opened irrespective of the membrane potential. Inactivation is mediated by the

25 channel's N-terminus, which acts as a plug that closes the pore. The transition from an inactive to a closed state requires a return to resting potential.

Voltage-gated  $\text{Na}^+$  channels are heterotrimeric complexes composed of a 260 kDa pore-forming  $\alpha$  subunit that associates with two smaller auxiliary subunits,  $\beta 1$  and  $\beta 2$ . The  $\beta 2$  subunit is an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with  $\alpha$

30 and  $\beta 1$  subunits correlates with increased functional expression of the channel, a change in its gating properties, as well as an increase in whole cell capacitance due to an increase in membrane surface area (Isom, L.L. et al. (1995) Cell 83:433-442).

Non voltage-gated  $\text{Na}^+$  channels include the members of the amiloride-sensitive  $\text{Na}^+$  channel/degenerin (NaC/DEG) family. Channel subunits of this family are thought to consist of two

35 transmembrane domains flanking a long extracellular loop, with the amino and carboxyl termini

located within the cell. The NaC/DEG family includes the epithelial Na<sup>+</sup> channel (ENaC) involved in Na<sup>+</sup> reabsorption in epithelia including the airway, distal colon, cortical collecting duct of the kidney, and exocrine duct glands. Mutations in ENaC result in pseudohypoaldosteronism type 1 and Liddle's syndrome (pseudohyperaldosteronism). The NaC/DEG family also includes the recently characterized H<sup>+</sup>-gated cation channels or acid-sensing ion channels (ASIC). ASIC subunits are expressed in the brain and form heteromultimeric Na<sup>+</sup>-permeable channels. These channels require acid pH fluctuations for activation. ASIC subunits show homology to the degenerins, a family of mechanically-gated channels originally isolated from *C. elegans*. Mutations in the degenerins cause neurodegeneration. ASIC subunits may also have a role in neuronal function, or in pain perception, since tissue acidosis causes pain (Waldmann, R. and M. Lazdunski (1998) *Curr. Opin. Neurobiol.* 8:418-424; Eglen, R.M. et al. (1999) *Trends Pharmacol. Sci.* 20:337-342).

K<sup>+</sup> channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as Ca<sup>2+</sup> and cAMP. In non-excitable tissue, K<sup>+</sup> channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K<sup>+</sup> channels are responsible for setting the resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell contains a Na<sup>+</sup>-K<sup>+</sup> pump and ion channels that provide the redistribution of Na<sup>+</sup>, K<sup>+</sup>, and Cl<sup>-</sup>. The pump actively transports Na<sup>+</sup> out of the cell and K<sup>+</sup> into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K<sup>+</sup> and Cl<sup>-</sup> to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl<sup>-</sup> flows out of the cell. The flow of K<sup>+</sup> is balanced by an electromotive force pulling K<sup>+</sup> into the cell, and a K<sup>+</sup> concentration gradient pushing K<sup>+</sup> out of the cell. Thus, the resting membrane potential is primarily regulated by K<sup>+</sup> flow (Salkoff, L. and T. Jegla (1995) *Neuron* 15:489-492).

Potassium channel subunits of the Shaker-like superfamily all have the characteristic six transmembrane/1 pore domain structure. Four subunits combine as homo- or heterotetramers to form functional K channels. These pore-forming subunits also associate with various cytoplasmic  $\beta$  subunits that alter channel inactivation kinetics. The Shaker-like channel family includes the voltage-gated K<sup>+</sup> channels as well as the delayed rectifier type channels such as the human ether-a-go-go related gene (HERG) associated with long QT, a cardiac dysrhythmia syndrome (Curran, M.E. (1998) *Curr. Opin. Biotechnol.* 9:565-572; Kaczorowski, G.J. and M.L. Garcia (1999) *Curr. Opin. Chem. Biol.* 3:448-458).

A second superfamily of K<sup>+</sup> channels is composed of the inward rectifying channels (Kir). Kir channels have the property of preferentially conducting K<sup>+</sup> currents in the inward direction.

These proteins consist of a single potassium selective pore domain and two transmembrane domains, which correspond to the fifth and sixth transmembrane domains of voltage-gated K<sup>+</sup> channels. Kir subunits also associate as tetramers. The Kir family includes ROMK1, mutations in which lead to Bartter syndrome, a renal tubular disorder. Kir channels are also involved in  
5 regulation of cardiac pacemaker activity, seizures and epilepsy, and insulin regulation (Doupnik, C.A. et al. (1995) *Curr. Opin. Neurobiol.* 5:268-277; Curran, *supra*).

The recently recognized TWIK K<sup>+</sup> channel family includes the mammalian TWIK-1, TREK-1 and TASK proteins. Members of this family possess an overall structure with four transmembrane domains and two P domains. These proteins are probably involved in controlling  
10 the resting potential in a large set of cell types (Duprat, F. et al. (1997) *EMBO J* 16:5464-5471).

The voltage-gated Ca<sup>2+</sup> channels have been classified into several subtypes based upon their electrophysiological and pharmacological characteristics. L-type Ca<sup>2+</sup> channels are predominantly expressed in heart and skeletal muscle where they play an essential role in excitation-contraction coupling. T-type channels are important for cardiac pacemaker activity, while N-type and P/Q-type  
15 channels are involved in the control of neurotransmitter release in the central and peripheral nervous system. The L-type and N-type voltage-gated Ca<sup>2+</sup> channels have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The  $\alpha_1$  subunit forms the membrane pore and voltage sensor, while the  $\alpha_2\delta$  and  $\beta$  subunits modulate the voltage-dependence, gating properties, and the current amplitude of the  
20 channel. These subunits are encoded by at least six  $\alpha_1$ , one  $\alpha_2\delta$ , and four  $\beta$  genes. A fourth subunit,  $\gamma$ , has been identified in skeletal muscle (Walker, D. et al. (1998) *J. Biol. Chem.* 273:2361-2367; McCleskey, E.W. (1994) *Curr. Opin. Neurobiol.* 4:304-312).

The high-voltage-activated Ca<sup>2+</sup> channels that have been characterized biochemically include complexes of a pore-forming  $\alpha_1$  subunit of approximately 190-250 kDa; a  
25 transmembrane complex of  $\alpha_2$  and  $\delta$  subunits; an intracellular  $\beta$  subunit; and in some cases a transmembrane  $\gamma$  subunit. A variety of  $\alpha_1$  subunits,  $\alpha_2\delta$  complexes,  $\beta$  subunits, and  $\gamma$  subunits are known. The Cav1 family of  $\alpha_1$  subunits conduct L-type Ca<sup>2+</sup> currents, which initiate muscle contraction, endocrine secretion, and gene transcription, and are regulated primarily by second messenger-activated protein phosphorylation pathways. The Cav2  
30 family of  $\alpha_1$  subunits conduct N-type, P/Q-type, and R-type Ca<sup>2+</sup> currents, which initiate rapid synaptic transmission and are regulated primarily by direct interaction with G proteins and SNARE proteins and secondarily by protein phosphorylation. The Cav3 family of  $\alpha_1$  subunits conduct T-type Ca<sup>2+</sup> currents, which are activated and inactivated more rapidly and at more negative membrane potentials than other Ca<sup>2+</sup> current types. The distinct structures and patterns of  
35 regulation of these three families of Ca<sup>2+</sup> channels provide an array of Ca<sup>2+</sup> entry pathways in



response to changes in membrane potential and a range of possibilities for regulation of  $\text{Ca}^{2+}$  entry by second messenger pathways and interacting proteins (Catterall, W.A. (2000) *Annu. Rev. Cell Dev. Biol.* 16:521-555).

The  $\alpha$ -2 subunit of the voltage-gated  $\text{Ca}^{2+}$ -channel may include one or more Cache domains. An extracellular Cache domain may be fused to an intracellular catalytic domain, such as the histidine kinase, PP2C phosphatase, GGDEF (a predicted diguanylate cyclase), HD-GYP (a predicted phosphodiesterase) or adenylyl cyclase domain, or to a noncatalytic domain, like the methyl-accepting, DNA-binding winged helix-turn-helix, GAF, PAS or HAMP (a domain found in histidine kinases, adenylyl cyclases, ethyl-binding proteins and phosphatases). Small molecules are bound via the Cache domain and this signal is converted into diverse outputs depending on the intracellular domains (Anantharaman, V. and Aravind, L.(2000) *Trends Biochem. Sci.* 25:535-537).

The transient receptor family (Trp) of calcium ion channels are thought to mediate capacitative calcium entry (CCE). CCE is the  $\text{Ca}^{2+}$  influx into cells to resupply  $\text{Ca}^{2+}$  stores depleted by the action of inositol triphosphate (IP3) and other agents in response to numerous hormones and growth factors. Trp and Trp-like were first cloned from *Drosophila* and have similarity to voltage gated  $\text{Ca}^{2+}$  channels in the S3 through S6 regions. This suggests that Trp and/or related proteins may form mammalian CCE channels (Zhu, X. et al. (1996) *Cell* 85:661-671; Boulay, G. et al. (1997) *J. Biol. Chem.* 272:29672-29680). Melastatin is a gene isolated in both the mouse and human, whose expression in melanoma cells is inversely correlated with melanoma aggressiveness *in vivo*. The human cDNA transcript corresponds to a 1533-amino acid protein having homology to members of the Trp family. It has been proposed that the combined use of malastatin mRNA expression status and tumor thickness might allow for the determination of subgroups of patients at both low and high risk for developing metastatic disease (Duncan, L.M. et al (2001) *J. Clin. Oncol.* 19:568-576).

Chloride channels are necessary in endocrine secretion and in regulation of cytosolic and organelle pH. In secretory epithelial cells,  $\text{Cl}^-$  enters the cell across a basolateral membrane through an  $\text{Na}^+$ ,  $\text{K}^+/\text{Cl}^-$  cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of  $\text{Cl}^-$  from the apical surface, in response to hormonal stimulation, leads to flow of  $\text{Na}^+$  and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal genetic disorder in humans. CFTR is a member of the ABC transporter family, and is composed of two domains each consisting of six transmembrane domains followed by a nucleotide-binding site. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus", and

devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) *J. Exp. Biol.* 172:245-266).

The voltage-gated chloride channels (CLC) are characterized by 10-12 transmembrane domains, as well as two small globular domains known as CBS domains. The CLC subunits probably function as homotetramers. CLC proteins are involved in regulation of cell volume, membrane potential stabilization, signal transduction, and transepithelial transport. Mutations in CLC-1, expressed predominantly in skeletal muscle, are responsible for autosomal recessive generalized myotonia and autosomal dominant myotonia congenita, while mutations in the kidney channel CLC-5 lead to kidney stones (Jentsch, T.J. (1996) *Curr. Opin. Neurobiol.* 6:303-310).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of  $\text{Na}^+$  and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as  $\gamma$ -aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential. Neurotransmitter-gated ion channels have four transmembrane domains and probably function as pentamers (Jentsch, *supra*). Amino acids in the second transmembrane domain appear to be important in determining channel permeation and selectivity (Sather, W.A. et al. (1994) *Curr. Opin. Neurobiol.* 4:313-323).

Ligand-gated channels can be regulated by intracellular second messengers. For example, calcium-activated  $\text{K}^+$  channels are gated by internal calcium ions. In nerve cells, an influx of calcium during depolarization opens  $\text{K}^+$  channels to modulate the magnitude of the action potential (Ishi et al., *supra*). The large conductance (BK) channel has been purified from brain and its subunit composition determined. The  $\alpha$  subunit of the BK channel has seven rather than six transmembrane domains in contrast to voltage-gated  $\text{K}^+$  channels. The extra transmembrane domain is located at the subunit N-terminus. A 28-amino-acid stretch in the C-terminal region of the subunit (the "calcium bowl" region) contains many negatively charged residues and is thought to be the region responsible for calcium binding. The  $\beta$  subunit consists of two transmembrane domains connected by a glycosylated extracellular loop, with intracellular N- and C-termini (Kaczorowski, *supra*; Vergara, C. et al. (1998) *Curr. Opin. Neurobiol.* 8:321-329).

Cyclic nucleotide-gated (CNG) channels are gated by cytosolic cyclic nucleotides. The best examples of these are the cAMP-gated  $\text{Na}^+$  channels involved in olfaction and the cGMP-gated

cation channels involved in vision. Both systems involve ligand-mediated activation of a G-protein coupled receptor which then alters the level of cyclic nucleotide within the cell. CNG channels also represent a major pathway for  $\text{Ca}^{2+}$  entry into neurons, and play roles in neuronal development and plasticity. CNG channels are tetramers containing at least two types of subunits, an  $\alpha$  subunit which

5 can form functional homomeric channels, and a  $\beta$  subunit, which modulates the channel properties. All CNG subunits have six transmembrane domains and a pore forming region between the fifth and sixth transmembrane domains, similar to voltage-gated  $\text{K}^+$  channels. A large C-terminal domain contains a cyclic nucleotide binding domain, while the N-terminal domain confers variation among channel subtypes (Zufall, F. et al. (1997) *Curr. Opin. Neurobiol.* 7:404-412).

10 The activity of other types of ion channel proteins may also be modulated by a variety of intracellular signalling proteins. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Kir channels are activated by the binding of the  $\text{G}\beta\gamma$  subunits of heterotrimeric G-proteins (Reimann, F. and F.M. Ashcroft (1999) *Curr. Opin. Cell.*

15 *Biol.* 11:503-508). Other proteins are involved in the localization of ion channels to specific sites in the cell membrane. Such proteins include the PDZ domain proteins known as MAGUKs (membrane-associated guanylate kinases) which regulate the clustering of ion channels at neuronal synapses (Craven, S.E. and D.S. Brecht (1998) *Cell* 93:495-498).

#### Disease Correlation

20 The etiology of numerous human diseases and disorders can be attributed to defects in the transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g., cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small

25 molecules across membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) *Exp. Nephrol.* 4:253-262; Talente, G.M. et al. (1994) *Ann. Intern. Med.* 120:218-226; and Chillon, M. et al. (1995) *New Engl. J. Med.* 332:1475-1480).

Human diseases caused by mutations in ion channel genes include disorders of skeletal muscle, cardiac muscle, and the central nervous system. Mutations in the pore-forming subunits of

30 sodium and chloride channels cause myotonia, a muscle disorder in which relaxation after voluntary contraction is delayed. Sodium channel myotonias have been treated with channel blockers. Mutations in muscle sodium and calcium channels cause forms of periodic paralysis, while mutations in the sarcoplasmic calcium release channel, T-tubule calcium channel, and muscle sodium channel cause malignant hyperthermia. Cardiac arrhythmia disorders such as the long QT

35 syndromes and idiopathic ventricular fibrillation are caused by mutations in potassium and sodium

channels (Cooper, E.C. and L.Y. Jan (1998) Proc. Natl. Acad. Sci. USA 96:4759-4766). All four known human idiopathic epilepsy genes code for ion channel proteins (Berkovic, S.F. and L.E. Scheffer (1999) Curr. Opin. Neurology 12:177-182). Other neurological disorders such as ataxias, hemiplegic migraine and hereditary deafness can also result from mutations in ion channel genes

5 (Jen, J. (1999) Curr. Opin. Neurobiol. 9:274-280; Cooper, supra).

Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia.

Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol.

10 39:47-98). Various classes of ion channels also play an important role in the perception of pain, and thus are potential targets for new analgesics. These include the vanilloid-gated ion channels, which are activated by the vanilloid capsaicin, as well as by noxious heat. Local anesthetics such as lidocaine and mexiletine which blockade voltage-gated Na<sup>+</sup> channels have been useful in the treatment of neuropathic pain (Eglen, supra). Calcium-channel protein expression is altered in  
15 metastatic melanomas (Enklaar, T. et al. (2000) Genomics 67:179-187).

Ion channels in the immune system have recently been suggested as targets for immunomodulation. T-cell activation depends upon calcium signaling, and a diverse set of T-cell specific ion channels has been characterized that affect this signaling process. Channel blocking agents can inhibit secretion of lymphokines, cell proliferation, and killing of target cells. A peptide  
20 antagonist of the T-cell potassium channel Kv1.3 was found to suppress delayed-type hypersensitivity and allogenic responses in pigs, validating the idea of channel blockers as safe and efficacious immunosuppressants (Cahalan, M.D. and K.G. Chandy (1997) Curr. Opin. Biotechnol. 8:749-756).

The discovery of new transporters and ion channels, and the polynucleotides encoding  
25 them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of transport, neurological, muscle, immunological and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

## 30 SUMMARY OF THE INVENTION

The invention features purified polypeptides, transporters and ion channels, referred to collectively as "TRICH" and individually as "TRICH-1," "TRICH-2," "TRICH-3," "TRICH-4," "TRICH-5," "TRICH-6," "TRICH-7," "TRICH-8," "TRICH-9," "TRICH-10," "TRICH-11," "TRICH-12," "TRICH-13," "TRICH-14," "TRICH-15," "TRICH-16," "TRICH-17," "TRICH-18,"  
35 "TRICH-19," "TRICH-20," "TRICH-21," "TRICH-22," "TRICH-23," "TRICH-24," "TRICH-25,"

"TRICH-26," "TRICH-27," "TRICH-28," "TRICH-29," "TRICH-30," "TRICH-31," and "TRICH-32." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-32.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-32. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:33-64.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of

SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

The method comprises a) culturing a cell under conditions suitable for expression of the

- 5 polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid

- 10 sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group  
15 consisting of SEQ ID NO:1-32.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the  
20 group consisting of SEQ ID NO:33-64, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group  
25 consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-

- 30 d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if  
35 present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous

nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of  
5 SEQ ID NO:33-64, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain  
10 reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a  
15 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and a pharmaceutically acceptable excipient. In one embodiment,  
20 the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an  
25 agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an  
30 immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of  
35 treating a disease or condition associated with decreased expression of functional TRICH,

comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the



absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

10 The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, ii) 15 a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide 20 in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide 25 complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization 30 complex in the treated biological sample is indicative of toxicity of the test compound.

#### BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

35 Table 2 shows the GenBank identification number and annotation of the nearest GenBank

homolog for polypeptides of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

## DEFINITIONS

"TRICH" refers to the amino acid sequences of substantially purified TRICH obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

5 The term "agonist" refers to a molecule which intensifies or mimics the biological activity of TRICH. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

10 An "allelic variant" is an alternative form of the gene encoding TRICH. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of  
15 nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding TRICH include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TRICH or a polypeptide with at least one functional characteristic of TRICH. Included within this definition  
20 are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TRICH, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding TRICH. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a  
25 functionally equivalent TRICH. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of TRICH is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar  
30 side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or  
35 synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally

occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence.

- 5 Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of TRICH. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind TRICH polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an in vitro evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH<sub>2</sub>), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules,

e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

The term "intramer" refers to an aptamer which is expressed in vivo. For example, a  
5 vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed  
10 nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as  
15 phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring  
20 nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or  
25 "immunogenic" refers to the capability of the natural, recombinant, or synthetic TRICH, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement,  
30 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding TRICH or  
35 fragments of TRICH may be employed as hybridization probes. The probes may be stored in freeze-

dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

- 5 "Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, 10 Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows 15 amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
20	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
25	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
30	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
35	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical 40 conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the

absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which  
5 retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

10 "Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an  
15 exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of TRICH or the polynucleotide encoding TRICH which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise  
20 up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a  
25 molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

30 A fragment of SEQ ID NO:33-64 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:33-64, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:33-64 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:33-64 from related polynucleotide sequences. The precise length of a fragment of SEQ  
35 ID NO:33-64 and the region of SEQ ID NO:33-64 to which the fragment corresponds are routinely

determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-32 is encoded by a fragment of SEQ ID NO:33-64. A fragment of SEQ ID NO:1-32 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-32. For example, a fragment of SEQ ID NO:1-32 is useful as an

- 5 immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-32. The precise length of a fragment of SEQ ID NO:1-32 and the region of SEQ ID NO:1-32 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

- A "full length" polynucleotide sequence is one containing at least a translation initiation  
10 codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

- The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer  
15 to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

- Percent identity between polynucleotide sequences may be determined using the default  
20 parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters  
25 are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

- Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local  
30 Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST  
35 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2



Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool

- 5 Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Reward for match: 1*

*Penalty for mismatch: -2*

- 10 *Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

*Filter: on*

- 15 Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment  
20 length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

- Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic  
25 acid sequences that all encode substantially the same protein.

- The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative  
30 substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

- Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of  
35 polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap

penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Open Gap: 11 and Extension Gap: 1 penalties*

10 *Gap x drop-off: 50*

*Expect: 10*

*Word Size: 3*

*Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas

wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

- 5        Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for  
10        calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

- High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS,  
15        for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under  
20        particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

- 25        The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g.,  $C_0t$  or  $R_0t$  analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate  
30        substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

- "Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression  
35        of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect

cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TRICH which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide  
5 fragment of TRICH which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other  
10 chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of TRICH. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of TRICH.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide,  
15 oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a  
20 functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which  
25 comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an TRICH may involve lipidation, glycosylation,  
30 phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TRICH.

"Probe" refers to nucleic acid sequences encoding TRICH, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are  
35 isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule.

Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and

5 identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences.

10 Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold  
15 Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research,

20 Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer  
25 selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for  
30 Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human

35 Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple

sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing TRICH, nucleic acids encoding TRICH, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or

cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by

infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), *supra*.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, or at least 98%, or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, or at least 98%, or greater sequence identity over a certain defined length of one of the polypeptides.

## THE INVENTION



The invention is based on the discovery of new human transporters and ion channels (TRICH), the polynucleotides encoding TRICH, and the use of these compositions for the diagnosis, treatment, or prevention of transport, neurological, muscle, immunological and cell proliferative disorders.

5 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide  
10 sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the  
15 polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank homolog(s) along with relevant  
20 citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential  
25 phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

30 Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are transporters and ion channels. For example, SEQ ID NO:1 is 83% identical to human sodium-hydrogen exchanger 6 (GenBank ID g2944233) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.1e-242, which indicates the probability of obtaining the observed  
35 polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a sodium/hydrogen

exchanger family domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, and other BLAST analyses provide further corroborative evidence that SEQ ID NO:1 is a sodium-hydrogen exchange transporter. In another example, SEQ ID NO:7 is 85% identical to *Rattus norvegicus* Na<sup>+</sup>/K<sup>+</sup>-ATPase alpha subunit (GenBank ID g619915) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:7 also contains an E1-E2 ATPase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:7 is a cation-transporting ATPase. In yet another example, SEQ ID NO:13 is 77% identical to a human carrier-like protein (GenBank ID g3694661) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.5e-209, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:13 also contains a mitochondrial energy transfer protein domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) The presence of this domain is confirmed by BLIMPS, MOTIFS, and PROFILESCAN analyses, providing further corroborative evidence that SEQ ID NO:13 is a transporter. Further, SEQ ID NO:16 is 41% identical to human novel ATPase (GenBank ID g8979801) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.1e-165, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:16 also contains an E1-E2 ATPase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:16 is a cation-transporting ATPase. In a further example, SEQ ID NO:19 is 43% identical to *Sinorhizobium* sp. As4 ArsA, the catalytic subunit of the arsenic oxyanion-translocating ATPase (GenBank ID g5802945) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 7.7e-125, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:19 also contains an anion-transporting ATPase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from additional BLAST and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:19 is an anion-transporting ATPase. In yet a further

example, SEQ ID NO:21 is 54% identical to a murine putative E1-E2 ATPase (GenBank ID g28577) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $5.2e-190$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:21 also contains six transmembrane domains as determined using TMAP, a program which delineates transmembrane segments. (See Table 3.) Data from BLIMPS, and MOTIFS, analyses provide further corroborative evidence that SEQ ID NO:21 is an ATPase. In a further example, SEQ ID NO:24 is 52% identical, from residue A77 to residue L1007, to rat NMDAR-L (GenBank ID g2160125) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $1.5e-262$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:24 also contains a ligand gated ion channel domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS analysis provide further corroborative evidence that SEQ ID NO:24 is a glutamate receptor. SEQ ID NO:2-6, SEQ ID NO:8-12, SEQ ID NO:14-15, SEQ ID NO:17-18, SEQ ID NO:20, SEQ ID NO:22-23, and SEQ ID NO:25-32 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-32 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide sequences of the invention, and of fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:33-64 or that distinguish between SEQ ID NO:33-64 and related polynucleotide sequences.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences

including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL\_XXXXXX\_N<sub>1</sub>N<sub>2</sub>YYYYY\_N<sub>3</sub>N<sub>4</sub> represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is the number of the prediction generated by the algorithm, and N<sub>1,2,3,4</sub>, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FLXXXXXX\_gAAAAA\_gBBBBB\_1\_N is a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide

sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

- 5           The invention also encompasses TRICH variants. A preferred TRICH variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the TRICH amino acid sequence, and which contains at least one functional or structural characteristic of TRICH.

- 10           The invention also encompasses polynucleotides which encode TRICH. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:33-64, which encodes TRICH. The polynucleotide sequences of SEQ ID NO:33-64, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

- 15           The invention also encompasses a variant of a polynucleotide sequence encoding TRICH. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TRICH. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID  
20 NO:33-64 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:33-64. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TRICH.

- 25           In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding TRICH. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding TRICH, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50%  
30 polynucleotide sequence identity to the polynucleotide sequence encoding TRICH over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding TRICH. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or  
35 structural characteristic of TRICH.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding TRICH, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide  
5 sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring TRICH, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode TRICH and its variants are generally capable  
10 of hybridizing to the nucleotide sequence of the naturally occurring TRICH under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TRICH or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with  
15 which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TRICH and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode TRICH and  
20 TRICH derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding TRICH or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of  
25 hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:33-64 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

30 Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the  
35 ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence

preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding TRICH may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Appl. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Appl. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode TRICH may be cloned in recombinant DNA molecules that direct expression of TRICH, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express TRICH.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TRICH-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of TRICH, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with



fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding TRICH may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, TRICH itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of TRICH, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active TRICH, the nucleotide sequences encoding TRICH or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding TRICH. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TRICH. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TRICH and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ.

20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding TRICH and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding TRICH. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding TRICH. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding TRICH can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TRICH into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M.

Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

5 Yeast expression systems may be used for production of TRICH. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel,  
10 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of TRICH. Transcription of sequences encoding TRICH may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO  
15 J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or  
pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology  
20 (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TRICH may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to  
25 obtain infective virus which expresses TRICH in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of  
30 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression  
35 of TRICH in cell lines is preferred. For example, sequences encoding TRICH can be transformed

into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding TRICH is inserted within a marker gene sequence, transformed cells containing sequences encoding TRICH can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TRICH under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding TRICH and that express TRICH may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of TRICH using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing 5 monoclonal antibodies reactive to two non-interfering epitopes on TRICH is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, 10 Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding TRICH include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled 15 nucleotide. Alternatively, the sequences encoding TRICH, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by 20 Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding TRICH may be cultured under 25 conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode TRICH may be designed to contain signal sequences which direct secretion of TRICH through a prokaryotic or eukaryotic cell membrane.

30 In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. 35 Different host cells which have specific cellular machinery and characteristic mechanisms for

post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TRICH may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric TRICH protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TRICH activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TRICH encoding sequence and the heterologous protein sequence, so that TRICH may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled TRICH may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

TRICH of the present invention or fragments thereof may be used to screen for compounds that specifically bind to TRICH. At least one and up to a plurality of test compounds may be screened for specific binding to TRICH. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of TRICH, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which TRICH binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case,

the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express TRICH, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing TRICH or cell membrane fractions which contain TRICH are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TRICH or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with TRICH, either in solution or affixed to a solid support, and detecting the binding of TRICH to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

TRICH of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of TRICH. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for TRICH activity, wherein TRICH is combined with at least one test compound, and the activity of TRICH in the presence of a test compound is compared with the activity of TRICH in the absence of the test compound. A change in the activity of TRICH in the presence of the test compound is indicative of a compound that modulates the activity of TRICH. Alternatively, a test compound is combined with an in vitro or cell-free system comprising TRICH under conditions suitable for TRICH activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of TRICH may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding TRICH or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or

developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding TRICH may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding TRICH can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding TRICH is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TRICH, e.g., by secreting TRICH in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

## THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TRICH and transporters and ion channels. In addition, examples of tissues expressing TRICH can be found in Table 6. Therefore, TRICH appears to play a role in transport, neurological, muscle, immunological and cell proliferative disorders. In the treatment of disorders associated with increased TRICH expression or activity, it is desirable to decrease the expression or activity of TRICH. In the treatment of disorders associated with decreased TRICH expression or activity, it is desirable to increase the expression or activity of TRICH.

Therefore, in one embodiment, TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug



resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy,

5 mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia,

10 Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, glycogen storage disease, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, pseudohypoaldosteronism type 1, Liddle's syndrome, cystinuria, iminoglycinuria, Hartup

15 disease, Fanconi disease, and Bartter syndrome; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating

20 diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis,

25 encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic

30 paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, hemiplegic migraine, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular

35 dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy,

lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, acid maltase deficiency (AMD, also known as Pompe's disease), generalized myotonia, and myotonia congenita; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified TRICH in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of TRICH may be

administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those listed above.

In a further embodiment, an antagonist of TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH. Examples of such disorders include, but are not limited to, those transport, neurological, muscle, immunological and cell proliferative disorders described above. In one aspect, an antibody which specifically binds TRICH may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express TRICH.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of TRICH may be produced using methods which are generally known in the art. In particular, purified TRICH may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TRICH. Antibodies to TRICH may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with TRICH or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TRICH have an amino acid sequence consisting of at least about 5 amino acids, and generally will

consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of TRICH amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

- 5 Monoclonal antibodies to TRICH may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

- In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and  
15 Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TRICH-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-  
20 10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

- 25 Antibody fragments which contain specific binding sites for TRICH may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D.  
30 et al. (1989) *Science* 246:1275-1281.)

- Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between TRICH and its  
35 specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies

reactive to two non-interfering TRICH epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TRICH. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of TRICH-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple TRICH epitopes, represents the average affinity, or avidity, of the antibodies for TRICH. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular TRICH epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the TRICH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TRICH, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TRICH-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, *supra*, and Coligan et al. *supra*.)

In another embodiment of the invention, the polynucleotides encoding TRICH, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TRICH. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TRICH. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence

complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990)

5 Blood 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

10 In another embodiment of the invention, polynucleotides encoding TRICH may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase

15 (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature*

20 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida

25 albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in TRICH expression or regulation causes disease, the expression of TRICH from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in

30 TRICH are treated by constructing mammalian expression vectors encoding TRICH and introducing these vectors by mechanical means into TRICH-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev.*

35 *Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr.*

Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of TRICH include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). TRICH may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TRICH from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TRICH expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TRICH under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant")

discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding TRICH to cells which have one or more genetic abnormalities with respect to the expression of TRICH. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TRICH to target cells which have one or more genetic abnormalities with respect to the expression of TRICH. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TRICH to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to



those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TRICH to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for TRICH into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TRICH-coding RNAs and the synthesis of high levels of TRICH in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of TRICH into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding TRICH.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by

scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding TRICH. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding TRICH.

Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased TRICH expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding TRICH may be therapeutically useful, and in the treatment of disorders associated with decreased TRICH expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TRICH may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in

altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding TRICH is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding TRICH are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding TRICH. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable

excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of TRICH, antibodies to TRICH, and mimetics, agonists, antagonists, or inhibitors of

5 TRICH.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

10 Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar  
15 region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patto, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active  
20 ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising TRICH or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of  
25 the macromolecule. Alternatively, TRICH or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell  
30 culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example  
35 TRICH or fragments thereof, antibodies of TRICH, and agonists, antagonists or inhibitors of

TRICH, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu$ g to 100,000  $\mu$ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

#### DIAGNOSTICS

In another embodiment, antibodies which specifically bind TRICH may be used for the diagnosis of disorders characterized by expression of TRICH, or in assays to monitor patients being treated with TRICH or agonists, antagonists, or inhibitors of TRICH. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TRICH include methods which utilize the antibody and a label to detect TRICH in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring TRICH, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TRICH expression.

Normal or standard values for TRICH expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to TRICH under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of

- 5 TRICH expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding TRICH may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide  
10 sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of TRICH may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TRICH, and to monitor regulation of TRICH levels during therapeutic intervention.

- 15 In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TRICH or closely related molecules may be used to identify nucleic acid sequences which encode TRICH. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine  
20 whether the probe identifies only naturally occurring sequences encoding TRICH, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TRICH encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:33-64 or  
25 from genomic sequences including promoters, enhancers, and introns of the TRICH gene.

Means for producing specific hybridization probes for DNAs encoding TRICH include the cloning of polynucleotide sequences encoding TRICH or TRICH derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA  
30 polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding TRICH may be used for the diagnosis of disorders associated with expression of TRICH. Examples of such disorders include, but are not limited to, a  
35 transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic

- fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias,
- 5 peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g.,
- 10 Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., , neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease,
- 15 glucose-galactose malabsorption syndrome, glycogen storage disease, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, pseudohypoaldosteronism type 1, Liddle's syndrome, cystinuria, iminoglycinuria, Hartup disease, Fanconi disease, and Bartter syndrome; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease,
- 20 Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system
- 25 disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic
- 30 nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, hemiplegic migraine, tardive
- 35 dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive

supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, acid maltase deficiency (AMD, also known as Pompe's disease), generalized myotonia, and myotonia congenita; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding TRICH may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered TRICH expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding TRICH may be useful in assays



that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TRICH may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding TRICH in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of TRICH, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding TRICH, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TRICH may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding TRICH, or a fragment of a polynucleotide complementary to the polynucleotide encoding TRICH, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent

conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of TRICH include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this

information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

5 In another embodiment, TRICH, fragments of TRICH, or antibodies specific for TRICH may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to  
10 generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be  
15 generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

20 Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of  
25 pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test  
30 compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression  
35 of these genes are used to normalize the rest of the expression data. The normalization procedure is

useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National  
5 Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the  
10 treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

15 Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance  
20 under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson,  
25 *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are  
30 compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence  
35 data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for TRICH to quantify the levels of TRICH expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lucking, A. et al. (1999) Anal.

- 5 Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

- Toxicant signatures at the proteome level are also useful for toxicological screening, and  
10 should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult,  
15 due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

- In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of  
20 each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

- 25 In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount  
30 of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

- Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116;  
35 Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad.

Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

5 In another embodiment of the invention, nucleic acid sequences encoding TRICH may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during  
10 chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J.  
15 (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

20 Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding TRICH on a physical map and a specific disorder, or a predisposition to a specific disorder, may  
25 help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse,  
30 may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for  
35 further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide

sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, TRICH, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between TRICH and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TRICH, or fragments thereof, and washed. Bound TRICH is then detected by methods well known in the art. Purified TRICH can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TRICH specifically compete with a test compound for binding TRICH. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TRICH.

In additional embodiments, the nucleotide sequences which encode TRICH may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/254,303, U.S. Ser. No. 60/256,190, U.S. Ser. No. 60/257,504, U.S. Ser. No. 60/261,546, U.S. Ser. No. 60/262,832, U.S. Ser. No. 60/264,377, and U.S. Ser. No. 60/266,019, are expressly incorporated by reference herein.

## EXAMPLES

### I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD

database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids



were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

### III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.

Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from Homo sapiens, Rattus norvegicus, Mus musculus, Caenorhabditis elegans, Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Candida albicans (Incyte Genomics, Palo Alto CA); and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs,

stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:33-64. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

#### **IV. Identification and Editing of Coding Sequences from Genomic DNA**

Putative transporters and ion channels were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range

of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode transporters and ion channels, the encoded polypeptides were analyzed by querying against PFAM models for transporters and ion channels. Potential transporters and ion channels were also identified by homology to Incyte cDNA sequences that had been annotated as transporters and ion channels. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

#### **V. Assembly of Genomic Sequence Data with cDNA Sequence Data**

##### **"Stitched" Sequences**

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from

genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

#### "Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

#### **VI. Chromosomal Mapping of TRICH Encoding Polynucleotides**

The sequences which were used to assemble SEQ ID NO:33-64 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:33-64 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously

identified disease genes map within or in proximity to the intervals indicated above.

## VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which

- 5 RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

- Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of  
10 the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

15

- The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is  
20 calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100%  
25 identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

- Alternatively, polynucleotide sequences encoding TRICH are analyzed with respect to the  
30 tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female;  
35 genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous

system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding TRICH. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

#### 10 VIII. Extension of TRICH Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using  
15 OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one  
20 extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech),  
25 ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step  
30 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II  
35 (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the

concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, 5 digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector 10 (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase 15 (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were 20 diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides 25 designed for such extension, and an appropriate genomic library.

#### IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:33-64 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide 30 fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). 35 An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-

based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

#### X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

#### Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X



first strand buffer, 0.03 units/ $\mu$ l RNase inhibitor, 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, 40  $\mu$ M dCTP, 40  $\mu$ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> RNAs are synthesized by in vitro

- 5 transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction
- 10 samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

#### Microarray Preparation

- Sequences of the present invention are used to generate array elements. Each array element
- 15 is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

- 20 Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a
- 25 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

- 30 Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

- 35 Hybridization

Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

#### 10 Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores.

Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-

compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

#### **XI. Complementary Polynucleotides**

Sequences complementary to the TRICH-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TRICH. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of TRICH. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TRICH-encoding transcript.

#### **XII. Expression of TRICH**

Expression and purification of TRICH is achieved using bacterial or virus-based expression systems. For expression of TRICH in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TRICH upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRICH in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TRICH by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or

human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, TRICH is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from TRICH at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified TRICH obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII where applicable.

### XIII. Functional Assays

TRICH function is assessed by expressing the sequences encoding TRICH at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane

composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of TRICH on gene expression can be assessed using highly purified  
5 populations of cells transfected with sequences encoding TRICH and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by  
10 those of skill in the art. Expression of mRNA encoding TRICH and other genes of interest can be analyzed by northern analysis or microarray techniques.

#### **XIV. Production of TRICH Specific Antibodies**

TRICH substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used  
15 to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the TRICH amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are  
20 well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the  
25 oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-TRICH activity by, for example, binding the peptide or TRICH to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

#### **XV. Purification of Naturally Occurring TRICH Using Specific Antibodies**

30 Naturally occurring or recombinant TRICH is substantially purified by immunoaffinity chromatography using antibodies specific for TRICH. An immunoaffinity column is constructed by covalently coupling anti-TRICH antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

35 Media containing TRICH are passed over the immunoaffinity column, and the column is

washed under conditions that allow the preferential absorbance of TRICH (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TRICH binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TRICH is collected.

## 5 XVI. Identification of Molecules Which Interact with TRICH

Molecules which interact with TRICH may include transporter substrates, agonists or antagonists, modulatory proteins such as G $\beta$ y proteins (Reimann, *supra*) or proteins involved in TRICH localization or clustering such as MAGUKs (Craven, *supra*). TRICH, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TRICH, washed, and any wells with labeled TRICH complex are assayed. Data obtained using different concentrations of TRICH are used to calculate values for the number, affinity, and association of TRICH with the candidate molecules.

Alternatively, proteins that interact with TRICH are isolated using the yeast 2-hybrid system (Fields, S. and O. Song (1989) *Nature* 340:245-246). TRICH, or fragments thereof, are expressed as fusion proteins with the DNA binding domain of Gal4 or lexA, and potential interacting proteins are expressed as fusion proteins with an activation domain. Interactions between the TRICH fusion protein and the TRICH interacting proteins (fusion proteins with an activation domain) reconstitute a transactivation function that is observed by expression of a reporter gene. Yeast 2-hybrid systems are commercially available, and methods for use of the yeast 2-hybrid system with ion channel proteins are discussed in Niethammer, M. and M. Sheng (1998, *Meth. Enzymol.* 293:104-122).

TRICH may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Potential TRICH agonists or antagonists may be tested for activation or inhibition of TRICH ion channel activity using the assays described in section XVIII.

## XVII. Demonstration of TRICH Activity

Ion channel activity of TRICH is demonstrated using an electrophysiological assay for ion conductance. TRICH can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding TRICH. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A second plasmid which expresses any one of a number of marker genes, such as  $\beta$ -galactosidase, is co-transformed into the cells to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after

transformation under conditions appropriate for the cell line to allow expression and accumulation of TRICH and  $\beta$ -galactosidase.

Transformed cells expressing  $\beta$ -galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells  
5 are tested for differences in membrane conductance by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or  $\beta$ -galactosidase sequences alone, are used as controls and tested in parallel. Cells expressing TRICH will have higher anion or cation conductance relative to control cells. The contribution of TRICH to conductance can be confirmed by incubating the cells using antibodies specific for  
10 TRICH. The antibodies will bind to the extracellular side of TRICH, thereby blocking the pore in the ion channel, and the associated conductance.

Alternatively, ion channel activity of TRICH is measured as current flow across a TRICH-containing *Xenopus laevis* oocyte membrane using the two-electrode voltage-clamp technique (Ishi et al., *supra*; Jegla, T. and L. Salkoff (1997) *J. Neurosci.* 17:32-44). TRICH is subcloned into an  
15 appropriate *Xenopus* oocyte expression vector, such as pBF, and 0.5-5 ng of mRNA is injected into mature stage IV oocytes. Injected oocytes are incubated at 18 °C for 1-5 days. Inside-out macropatches are excised into an intracellular solution containing 116 mM K-gluconate, 4 mM KCl, and 10 mM Hepes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the TRICH mediator, such as cAMP, cGMP, or  $\text{Ca}^{+2}$  (in the form of  $\text{CaCl}_2$ ), where appropriate.  
20 Electrode resistance is set at 2-5 M $\Omega$  and electrodes are filled with the intracellular solution lacking mediator. Experiments are performed at room temperature from a holding potential of 0 mV. Voltage ramps (2.5 s) from -100 to 100 mV are acquired at a sampling frequency of 500 Hz. Current measured is proportional to the activity of TRICH in the assay. In particular, the activity of TRICH-1 is measured as  $\text{Na}^+$  conductance and the activity of TRICH-3 is measured as  $\text{Ca}^{2+}$   
25 conductance.

Transport activity of TRICH is assayed by measuring uptake of labeled substrates into *Xenopus laevis* oocytes. Oocytes at stages V and VI are injected with TRICH mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM  $\text{CaCl}_2$ , 1mM  $\text{MgCl}_2$ , 1mM  $\text{Na}_2\text{HPO}_4$ , 5 mM Hepes, 3.8 mM NaOH, 50 $\mu\text{g/ml}$  gentamycin, pH 7.8) to  
30 allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM  $\text{CaCl}_2$ , 1mM  $\text{MgCl}_2$ , 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with  $^3\text{H}$ , fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times  
35 in  $\text{Na}^+$ -free medium, measuring the incorporated label, and comparing with controls. TRICH

activity is proportional to the level of internalized labeled substrate. Test substrates include ran-GTP for TRICH-2, glucose for TRICH-5, amino acids for TRICH-6 and TRICH-14, cations for TRICH-7 and TRICH-16, Na<sup>+</sup>, K<sup>+</sup> and Cl<sup>-</sup> ions for TRICH-15, reduced folate or analogues such as methotrexate for TRICH-17, divalent cations for TRICH-18, anions such as arsenate and antimonite for TRICH-19, and nitrate or oligopeptides for TRICH-20.

ATPase activity associated with TRICH can be measured by hydrolysis of radiolabeled ATP-[ $\gamma$ -<sup>32</sup>P], separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered <sup>32</sup>P using a scintillation counter. The reaction mixture contains ATP-[ $\gamma$ -<sup>32</sup>P] and varying amounts of TRICH in a suitable buffer incubated at 37°C for a suitable period of time. The reaction is terminated by acid precipitation with trichloroacetic acid and then neutralized with base, and an aliquot of the reaction mixture is subjected to membrane or filter paper-based chromatography to separate the reaction products. The amount of <sup>32</sup>P liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of TRICH in the assay.

Lipocalin activity of TRICH is measured by ligand fluorescence enhancement spectrofluorometry (Lin et al. (1997) Molecular Vision 3:17). Examples of ligands include retinol (Sigma, St. Louis MO) and 16-anthyloxy-palmitic acid (16-AP) (Molecular Probes Inc., Eugene OR). Ligand is dissolved in 100% ethanol and its concentration is estimated using known extinction coefficients (retinol: 46,000 A/M/cm at 325 nm; 16-AP: 8,200 A/M/cm at 361 nm). A 700  $\mu$ l aliquot of 1  $\mu$ M TRICH in 10 mM Tris (pH 7.5), 2 mM EDTA, and 500 mM NaCl is placed in a 1 cm path length quartz cuvette and 1  $\mu$ l aliquots of ligand solution are added. Fluorescence is measured 100 seconds after each addition until readings are stable. Change in fluorescence per unit change in ligand concentration is proportional to TRICH activity.

#### XVIII. Identification of TRICH Agonists and Antagonists

TRICH is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293. Ion channel activity of the transformed cells is measured in the presence and absence of candidate agonists or antagonists. Ion channel activity is assayed using patch clamp methods well known in the art or as described in Example XVII. Alternatively, ion channel activity is assayed using fluorescent techniques that measure ion flux across the cell membrane (Veliccebi, G. et al. (1999) Meth. Enzymol. 294:20-47; West, M.R. and C.R. Molloy (1996) Anal. Biochem. 241:51-58). These assays may be adapted for high-throughput screening using microplates. Changes in internal ion concentration are measured using fluorescent dyes such as the Ca<sup>2+</sup> indicator Fluo-4 AM, sodium-sensitive dyes such as SBFI and sodium green, or the Cl<sup>-</sup> indicator MQAE (all available from Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In a more generic version of this assay, changes in



membrane potential caused by ionic flux across the plasma membrane are measured using oxonyl dyes such as DiBAC<sub>4</sub> (Molecular Probes). DiBAC<sub>4</sub> equilibrates between the extracellular solution and cellular sites according to the cellular membrane potential. The dye's fluorescence intensity is 20-fold greater when bound to hydrophobic intracellular sites, allowing detection of DiBAC<sub>4</sub> entry into the cell (Gonzalez, J.E. and P.A. Negulescu (1998) Curr. Opin. Biotechnol. 9:624-631).

5 Candidate agonists or antagonists may be selected from known ion channel agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

10

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
7484831	1	7484831CD1	33	7484831CB1
2477266	2	2477266CD1	34	2477266CB1
3552033	3	3552033CD1	35	3552033CB1
4778139	4	4778139CD1	36	4778139CB1
4787433	5	4787433CD1	37	4787433CB1
7483598	6	7483598CD1	38	7483598CB1
7484823	7	7484823CD1	39	7484823CB1
143935	8	143935CD1	40	143935CB1
5923789	9	5923789CD1	41	5923789CB1
6046484	10	6046484CD1	42	6046484CB1
7481427	11	7481427CD1	43	7481427CB1
7483595	12	7483595CD1	44	7483595CB1
3788427	13	3788427CD1	45	3788427CB1
6972455	14	6972455CD1	46	6972455CB1
8077668	15	8077668CD1	47	8077668CB1
55120485	16	55120485CD1	48	55120485CB1
3112883	17	3112883CD1	49	3112883CB1
4253888	18	4253888CD1	50	4253888CB1
7479974	19	7479974CD1	51	7479974CB1
7483850	20	7483850CD1	52	7483850CB1
5508353	21	5508353CD1	53	5508353CB1
8543628	22	8543628CD1	54	8543628CB1
7482754	23	7482754CD1	55	7482754CB1
3794818	24	3794818CD1	56	3794818CB1
4717525	25	4717525CD1	57	4717525CB1
5091793	26	5091793CD1	58	5091793CB1
5945527	27	5945527CD1	59	5945527CB1
6941124	28	6941124CD1	60	6941124CB1
6972530	29	6972530CD1	61	6972530CB1

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
6991750	30	6991750CD1	62	6991750CBI
71726948	31	71726948CD1	63	71726948CBI
7487393	32	7487393CD1	64	7487393CBI

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	7484831CD1	g2944233	5.1e-242	Sodium-hydrogen exchanger 6 [Homo sapiens] (Munata, M. et al. (1998) J. Biol. Chem. 273:6951-6959)
2	2477266CD1	g2102696	1.0e-37	[Homo sapiens] karyopherin beta 3 Yaseen, N.R. and Blobel, G. (1997) Proc. Natl. Acad. Sci. U.S.A. 94:4451-4456
3	3552033CD1	g3243075	0.0	[Homo sapiens] melastatin 1 Hunter, J.J. et al. (1998) Genomics 1998 54:116-123
4	4778139CD1	g8131903	5.1e-107	[Mus musculus] transient receptor potential-related protein
5	4787433CD1	g2337865	2.3e-251	[Homo sapiens] organic cation transporter
6	7483598CD1	g6978016	1.9e-32	[Rattus norvegicus] neuronal glutamine transporter Varoqui, H. et al. (2000) J. Biol. Chem. 275:4049-4054
7	7484823CD1	g619915	0.0	[Rattus norvegicus] Na,K-ATPase alpha subunit Shamraj, O.I., and Lingrel, J.B. (1994) Proc. Natl. Acad. Sci. USA 91:12952-12956
8	143935CD1	g179304	7.8e-116	Bi2 protein [Homo sapiens] (Wolf, F.W. et al. (1992) J. Biol. Chem. 267:1317-1326)
9	5923789CD1	g1552526	0.0	sodium-calcium exchanger form 3 [Rattus norvegicus] (Nicoll, D.A. et al. (1996) J. Biol. Chem. 271:24914-24921)
10	6046484CD1	g3243075	0.0	melastatin 1 [Homo sapiens] (Hunter, J.J. (1998) Genomics 54:116-123)
11	7481427CD1	g178661	9.8e-93	adenine nucleotide translocator-2 [Homo sapiens] (Ku, D.H. et al. (1990) J. Biol. Chem. 265: 16060-16063)

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
12	7483595CD1	g9453726	3.7e-61	ba48209.2 (Novel sulphate transporter family member) [Homo sapiens]
13	3788427CD1	g3694661	5.5e-209	carrier protein-like; similar to Q01888 (PID:g266574) [Homo sapiens]
14	6972455CD1	g4155688	1.7e-24	[Helicobacter pylori J99] AMINO ACID ABC TRANSPORTER, BINDING PROTEIN PRECURSOR
15	8077668CD1	g5081312	6.6e-47	[Rattus norvegicus] bumetanide-sensitive Na-K-2Cl cotransporter Anzai, N., et al. (1999) Roles of vasopressin and hypertonicity in basolateral Na/K/2Cl cotransporter expression in rat kidney inner medullary collecting duct cells. Jpn. J. Physiol. 49, 201-206
16	55120485CD1	g8979801	1.1e-165	cd37c10.3 (novel ATPase) [Homo sapiens]
17	3112883CD1	g3115983	4.0e-128	cd206d15.1 (Reduced Folate Carrier protein RFC LIKE) [Homo sapiens]
18	4253888CD1	g3925431	1.6e-29	[Caenorhabditis elegans] (Z82084) contains similarity to Pfam domain: PF01769 (Divalent cation transporter), Score=211.5, E-value=4.2e-60, N=2
19	7479974CD1	g5802945	7.7e-125	[Sinorhizobium sp. As4] Arsa (catalytic subunit of arsenic oxyanion-translocating ATPase)
20	7483850CD1	g11933414	6.3e-11	[Glycine max] nitrate transporter NRT1-5
21	5508353CD1	g6457270	5.2e-190	[Mus musculus] putative E1-E2 ATPase Halleck, M.S. et al., (1999) Physiol. Genomics (Online) 1:139-150 MEDLINE : 20473714
22	8543628CD1	g6967939	3.7e-45	[Campylobacter jejuni] amino-acid ABC transporter integral membrane protein Takamori S. et al., (2000) Nature 407:189-94

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
23	7482754CD1	g11640743	1.9e-20	[Homo sapiens] amino acid transporter system A1 Wang H. et al., (2000) Biochem. Biophys. Res. Commun. 273:1175-9
24	3794818CD1	g2160125	1.5e-262	NMDAR-L [Rattus norvegicus] Sucher, N.J. et al. (1995) J. Neurosci. 15:6509-6520
25	4717525CD1	g6841066	7.8e-111	calcium-binding transporter [Homo sapiens]
26	5091793CD1	g3880532	1.1e-51	Similarity to multidrug resistance protein (SW:BMRL_BACSU) [Caenorhabditis elegans] The C. elegans Sequencing Consortium (1998) Science 282: 2012-2018
27	5945527CD1	g7543982	1.5e-161	glycerol 3-phosphate permease [Homo sapiens]
28	6941124CD1	g476222	6.8e-66	anion exchanger 3 brain isoform [Homo sapiens] Yannoukakos, D. et al. (1994) Circ. Res. 75:603-614
29	6972530CD1	g10175963	3.5e-16	potassium channel protein [Bacillus halodurans] Takami, H. et al. (1999) Extremophiles 3:21-28
30	6991750CD1	g6273849	5.5e-11	cardiac sodium-calcium exchanger [Onchocerca mykiss] Xue, X.H. et al. (1999) Am. J. Physiol. 277:C693-C700
31	71726948CD1	g1628579	1.0e-152	sodium iodide symporter [Homo sapiens] Smanik, P.A. et al. (1996) Biochem. Biophys. Res. Commun. 226:339-345
32	7487393CD1	g7707622	2.2e-118	organic anion transporter 4 [Homo sapiens] Cha, S.H. et al. (2000) J. Biol. Chem. 275:4507-4512

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	7484831CD1	726	S11 S52 S66 S124 S147 S198 S244 S260 S546 S585 S689 S694 S695 S712 T59 T133 T154 T177 T591 T658 T665 T684	N145 N401 N572 N589 N674	Signal peptide: M1-A37	HMMER
					Transmembrane domains: P18-A39, R68-I88, R95-I115, K175-H203, L208-L236, D245-A269, A282-Q306, A319-L347, L360-L388, Y428-G456, H482-T502, Q508-L536 N-terminus is non-cytosolic	TMAP
					Sodium/hydrogen exchanger family: L74-V540	HMMER-PFAM
					Na <sup>+</sup> /H <sup>+</sup> exchanger isoform signatures PR01088: S44-A63, E64-I88, W89-I107, Y108-Q134, S299-D316, A318-M337, G588-D606, P612-Q640, V641-D668	BLIMPS-PRINTS
					Na <sup>+</sup> /H <sup>+</sup> exchanger signatures PR01084: V182-F193, G196-S210, I211-T219, G236-T266	BLIMPS-PRINTS
					Na <sup>+</sup> transport exchanger PD01672: V182-M230, A319-V344, F381-F414, F419-S465, F466-T512	BLIMPS-PRODOR
					Sodium hydrogen exchanger 6, myeloblast PD177855: Y557-N725, G527-E547	BLAST-PRODOR
					Na <sup>+</sup> /H <sup>+</sup> transmembrane transport antiporter, exchanger PD00631: E181-R539, L74-G125	BLAST-PRODOR
					Beta Na exchanger: DM02572 P48764 10-734: D171-R539, R21-L116 DM02572 Q01345 12-703: D179-D563, L22-R117 DM02572 P50482 16-723: E181-F558, G16-S124 DM02572 P26434 14-716: D179-D563, A48-T120, F605-P620	BLAST-DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2	2477266CD1	1081	S47 S79 S148 S180 S192 S315 S493 S615 S639 S697 S1011 T18 T25 T61 T147 T167 T328 T532 T586 T786 T813 T871 T881 T907 T974	N165 N686	Transmembrane domains: P197-L213 R255-V275 E498-P520 I839-V863 N terminus cytosolic	TMAP
					IMPORTIN SUBUNIT KARYOPHERIN PROTEIN TRANSPORT REPEAT PD014526: F691-D1033 PD014366: A458-S615, Q389-N412 Leucine zipper pattern L177-L198 Phospholipase A2 histidine active site C725-C732	BLAST_PRODOR
						MOTIFS
						MOTIFS
3	3552033CD1	1172	S212 S235 S300 S366 S401 S528 S558 S618 S687 S688 S884 S1017 S1059 S1060 S1069 S1076 S1088 S1125 T9 T147 T422 T459 T460 T917 T962 T984 T1031 T1112 T1118 T1132 T1155 Y645 Y857	N144 N233 N298 N420 N576 N579 N789 N915 N960 N1058 N1074	Transient receptor: Y943-M1001, R817-E882, E750-L807, D562-W608	HMMPER_PFAM
					Transmembrane domains: G51-I75 D397-R425 F712-V740 E786-R806 V822-G842 M852-A872 W934-T962 N terminus cytosolic	TMAP
					Transient receptor potential family signature PR01097: A941-T962, F963-F976, V990-M1003	BLIMPS_PRINTS
					PROTEIN MELASTATIN CHROMOSOME TRANSMEMBRANE PD018035: M1-L333 PD151509: I829-L1117 PD039592: E464-E660 PD022180: W328-R438	BLAST_PRODOR



Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3					ANK MOTIF REPEAT DM03196 P34586 38-822: I819-C1009, L583-G619, L702-L807, D114-N144, T9-Q48 Aminoacyl-transfer RNA synthetases class-II signature 2 A1111-L1120	BLAST_DOMO MOTIFS
4	4778139CD1	742	S42 S104 S135 S194 S203 S214 S220 S221 S234 S239 S277 S319 S339 S352 S354 S403 S435 S438 S479 S491 S510 S722 S735 T111 S725 T371 T508 T575 T619 T635 T731	N238 N258 N294 N650 N711	KINASE TRANSFERASE SERINE/THREONINE PROTEIN ATP BINDING ELONGATION FACTOR EEF2K CALCIUM/CALMODULINDEPENDENT EUKARYOTIC PD011701: K536-R709	BLAST_PRODOM
5	4787433CD1	577	S70 S119 S176 S319 S337 S544 S550 S560 T135 T356 T521 T534 T535 T569 Y10	N31 N57 N65 N68 N108 N345 N352 N546 N558	Sugar (and other) transporter: K120-E538	HMMER_PPFAM
					Transmembrane domains: G17-G45 R150-R178 L185-Y205 F214-I234 S243-L263 L269-F293 I355-S377 N390-D410 T416-P436 L442-Y462 A488-L516 N terminus cytosolic	TMAP
					Na+/H+ exchanger isoform PR010870 I32-V46 Leucine zipper pattern L146-L167	BLIMPS_PRINTS MOTIFS
6	7483598CD1	462	S24 S56 S90 S242 S243 S393 T282 T391	N100 N331 N436 N441 N457	Transmembrane amino acid transporter protein: S56-G412	HMMER_PPFAM
					Transmembrane domains: C33-L53 G65-K85 T102-A130 N148-L168 A175-V195 W213-Y241 A250-F278 F328-A348 L358-P378 T391-N419 N terminus non-cytosolic	TMAP
					ACID AMINO PROTEIN TRANSPORTER PERMEASE TRANSMEMBRANE INTERGENIC REGION PUTATIVE PROLINE PD001875: F39-T209	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
7	7484823CD1	1018	S10 S49 S151 S157 S372 S463 S525 S585 S648 S732 S938 T44 T82 T250 T344 T393 T399 T404 T444 T482 T618 T635 T755 T934 Y461 Y889	N212 N480	E1-E2 (cation transport) ATPase: V132-T363	HMMER_PPFAM
					Na+/K+ ATPase C-terminus: R829-Y1017, E30-S113	HMMER_PPFAM
					Transmembrane domains: H287-L315 E781-T809 I845-P873 V909-I931 A972-R1000 N terminus non-cytosolic	TMAP
					E1-E2 ATPases phosphoryl BL00154: V329-G365, T367-V385, K504-C514, D588-I628, V707-G730, G733-N766, G185-L202	BLIMPS_BLOCKS
					E1-E2 ATPases phosphorylation site atpase_e1_e2.prf: L354-E401	PROFILES CAN
					P-type cation-transporting atpase superfamily signature PR00119: D211-S225, C371-V385, G582-A593, A604-D614, T710-M729, S734-L746	BLIMPS_PRINTS
					H+-transporting ATPase (proton pump) signature PR00120: E682-E698, T710-G726, D742-L767	BLIMPS_PRINTS
					Sodium/potassium-transporting ATPase signature PR00121: L100-I114, L127-Q147, L291-G313, L364-V385, L501-L519, I782-L803, Y849-A869, F911-I931, R945-M969	BLIMPS_PRINTS
					ATPASE TRANSMEMBRANE TRANSPORT PUMP MAGNESIUM PD000132: V132-Y312, W314-N426, D667-E820, K473-D742, F118-S225, V581-I628	BLAST_PROD OM
					CALCIUM PD000121: L643-N749, I587-I629	
					CALCIUM PD000388: K828-Y1017	
					PTYPE TRANSPORTING ATPASE 1 PD111120: F750-H842	BLAST_PROD OM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
7					E1-E2 ATPASES PHOSPHORYLATION SITE DM00115 P50993 80-807: P79-D806 A34474 80-807: P79-D806 P06686 80-807: P79-D806 P24797 77-804: P79-D806 E1-E2 ATPases phosphorylation site D373-T379	BLAST_DOMO
8	143935CD1	313	S23 S30 S62 S101 S145 S146 S156 S176 S193 T51 T69 T235 T240	N166	K+ channel tetramerisation domain:K12-Q129	HMMER_PFAM
					Na+/H+ exchanger isoform PR01085H: T133-S145	BLIMPS_PRINTS
					EDP1 TNF ALPHA INDUCED ENDOTHELIAL B12 PD037429: L109-Q313	BLAST_PRODOM
					signal_cleavage: M1-T21	MOTIFS
9	5923789CD1	921	S69 S144 S151 S312 S381 S382 S691 S713 S720 S794 T106 T113 T125 T194 T267 T277 T460 T522 T572 T583 T594 T597 T632 T637 T672 Y405 Y608	N45 N130 N135 N817	Sodium/calcium exchanger protein: L757-L905, R110-F257	HMMER_PFAM
					PRECURSOR TRANSPORT SIGNAL GLYCOPROTEIN NA+/CA2+EXCHANGER SYMPORT TRANSMEMBRANE PD004181: W249-E390 PD001766: E385-H528 PD149743: V674-I777 PD149807: A529-G627	BLAST_PRODOM
					SODIUM/CALCIUM EXCHANGER CHAIN DM05297 P48768 1-920: C48-F921 DM05297 P48765 16-969: L6-E625, I628-F921 do ANTIPTORTER; MURZ; III; 34.7; DM02122 S20969 450-604: N130-W249	BLAST_DOMO
						BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9					Transmembrane domains: A2-R29, K73-S101, P167-P190, T194-M219, T237-M261, S720-P741, G748-T776, T776-D797, I815-W841, H849-R877, C891-T911 N-terminus is cytosolic	TMAP
					Signal cleavage: M1-A30	SPSCAN
					Signal Peptide: M1-A32	HMME
10	5046484CD1	1466	S86 S212 S235 S300 S366 S401 S518 S548 S608 S660 S719 S858 S988 S1011 S1039 S1040 S1049 S1056 S1087 S1114 S1166 S1224 S1234 S1323 S1399 S1410 T9 T422 T449 T450 T629 T637 T742 T941 T963 T1164 T1268 T1297 T1387 T1389 T1449 T1462 Y635	N133 N144 N233 N298 N420 N566 N569 N763 N1054 N1245	Transient receptor potential family PR01097: A920-T941, F942-F955 MELASTATIN 1 PD183973: L1097-C1466 MELASTATIN CHROMOSOME TRANSMEMBRANE C05C12.3 T01H8.5 I F54D1.5 IV PD018035: M1-L333 PD151509: I803-L1097 PD039592: E454-T652 Transmembrane domains: G51-I75, D397-R425, L602-Y630, F692-V712, W717-I737, N763-N783, Y789-F809, Y818-A846, W913-T941, L970-L987 N-terminus is cytosolic	HMME_PPFAM
11	7481427CD1	222	S28 S142 S159 T144		Mitochondrial carrier proteins: S7-K105, Y112-T202	HMME_PPFAM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
11					Mitochondrial energy transfer proteins BL00215: L13-Q37, I158-G170 Mitochondrial energy transfer proteins Mitoch_carrier.prf: C110-I158 Mitochondrial energy transfer proteins: P127-A136 Mitochondrial carrier protein signature PR00926: G120-D138, Y168-F186, D11-T24, T24-M38, G73-D93 Adenine nucleotide transfer protein PR00927: F8-A20, C51-R72, T84-K96, R111-V124, R146-L167, S207-S222 PROTEIN TRANSPORT TRANSMEMBRANE REPEAT MITOCHONDRION CARRIER MEMBRANE INNER MITOCHONDRIAL ADP/ATP PD000117: S7-A123, Y112-S222 MITOCHONDRIAL ENERGY TRANSFER PROTEINS DM00026 P02722 11-96: F12-I98 DM00026 S31935 14-108: F12-T107 DM00026 P02722 116-205: L117-N201 DM00026 S31935 110-208: Q108-N201 Transmembrane domains: I164-R187 N-terminus is cytosolic	BLIMPS_BLOCKS PROFILES SCAN MOTIFS BLIMPS_PRINTS BLIMPS_PRINTS BLAST_PRODROM BLAST_DOMO TMAP
12	7483595CD1	461	S168 S212 S233 S338 S362 S382 S439 T318 T348	N98 N163 N288 N344 N380 N381	SULFATE TRANSPORTER PROTEIN TRANSPORT TRANSMEMBRANE AFFINITY GLYCOPROTEIN SULPHATE HIGH DISEASE PD001755: R216-D272, V414-D459 SULFATE TRANSPORTERS DM01229 P40879 5-462: T23-A140, Q124-L222 DM01229 P45380 10-468: P130-L222, T23-L138 DM01229 P50443 49-505: I136-K247, T23-S144 Transmembrane domains: F6-P26, L31-S51, M62-S82, C91-A111, E128-K156, L161-R186 N-terminus is cytosolic Mitochondrial carrier proteins: S361-Y461, S266-Q359, T172-H264	BLAST_PRODROM BLAST_DOMO TMAP HMMER_PFAM
13	3788427CD1	502	S146 S304 S446 S467 T25 T59 T96 T104 T164 T385 T492 T501			

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13					Mitochondrial energy transfer proteins mitoch_carrier.prfl1: L362-I417 mitoch_carrier.prfl2: S270-I315 Mitochondrial energy transfer proteins: P287-I296 Mitochondrial carrier protein signature PR00926: G232-R252, V280-Q298, Y325-L343, V369-Q391 PROTEIN TRANSPORT TRANSMEMBRANE REPEAT MITOCHONDRIAL CARRIER MEMBRANE INNER MITOCHONDRIAL ADP/ATP PD000117: Y305-R453 MITOCHONDRIAL ENERGY TRANSFER PROTEINS DM00026 Q01888 126-214: G263-I352 DM00026 P29518 233-310: I271-I352 Transmembrane domains: R176-G204, A323-K351, L424-K452 N-terminus is cytosolic signal_cleavage: M1-A23	PROFILES SCAN MOTIFS BLIMPS_PRINTS BLAST_PRODROM
14	6972455CD1	261	S6 S26 S135 S198 T2 T36 T64 T84 T92 T165 Y218	N82 N172 N173	Signal Peptide: M1-P22, M1-A23, M1-A25, M1-G27 Bacterial extracellular solute-binding protein domain: M1-L253 Transmembrane domains: T175-A196 N-terminus is non-cytosolic Bacterial extracellular solute binding protein signature BL01039: G52-L72, R86-Y118, L90-S101 BACTERIAL EXTRACELLULAR SOLUTE-BINDING PROTEINS, FAMILY 3 DM00557 P27676 32-261: S19-K223 DM00557 P30860 5-241: L37-W252 DM00557 P39174 15-260: L8-W252 DM00557 P45678 11-258: A11-W252	SPSCAN HMHER HMHER_PPFAM TMAP BLIMPS_BLOCKS BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	8077668CD1	570	S40 S82 S265 S314 S349 S442 S446 S551 T50 T324 T344	N11 N342 N440	Transmembrane domains: V163-V183 N194-T214 V232-I259 I275-R302 S353-N373 A382-G402 P465-G481 N-terminus is cytosolic do SENSITIVE; COTRANSPORTER; SODIUM; CHLORIDE; DM01337 P55011 409-906: V278-G402 DM01337 P55013 381-879: V278-G402 DM01337 A53491 381-879: V278-G402 DM01337 P55014 297-795: V278-G402	TMAP BLAST_DOMO
16	55120485CD1	1033	S5 S103 S159 S241 S249 S338 S567 S587 S671 S798 S833 S850 S1008 S1028 T78 T97 T172 T375 T490 T664 T701 T784 T859 T871 Y85	N540 N669 N781 N819 N848 N867 N875 N1005	Transmembrane domains: F24-Y52 K197-L217 Y223-Y243 L394-Y422 D429-N454 T877-F893 T903-L931 L937-E964 N-terminus is non-cytosolic	TMAP
					E1-E2 ATPase domain: V268-P324 E1-E2 ATPases phosphorylation site signature BL00154: G284-L301, V442-G478, I480-L498, K624-C634, N695-M735 E1-E2 ATPases phosphorylation site: I466-F515	HMMER_Pfam BLIMPS_BLOCKS PROFILES SCAN
					P-type cation-transporting atpase superfamily signature PR00119: N309-T323, C484-L498, A711-D721 Sodium/potassium-transporting ATPase signature PR00121: C477-I498, V621-V639	BLIMPS_PRINTS BLIMPS_PRINTS
					ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATP BINDING TRANSPORT PUMP CALCIUM MAGNESIUM MEMBRANE PD000132: I230-D494	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
16					E1-E2 ATPASES PHOSPHORYLATION SITE DM00115 P22189 49-801: Q597-P738, W196-K322, P389-P447, S671-L723 DM00115 P37278 58-755: I192-I743, V822-K840 DM00115 P47317 26-695: F600-E749, I230-K559, K815-E853 DM00115 P54707 97-825: I232-I743, L813-E853	BLAST_DOMO
					E1-E2 ATPases phosphorylation site: D486-T492	MOTIFS
17	3112883CD1	496	S40 S211 S241 S251 S274 S275 S463 S466 S482 S493 T47 T117	N45 N166 N256	Reduced folate carrier domain: S10-V441	HMMEF_PPFAM
					Transmembrane domains: L8-M28 N53-Y76 Y79-Q107 F111-R138 L148-S168 V174-K194 K276-D304 N316-Y336 D342-L362 A367-A395 L405-V425 P434-L454 N-terminus is cytosolic	TMAP
					FOLATE CARRIER PROTEIN REDUCED TRANSPORTER GLYCOPROTEIN FOLATE BINDING TRANSPORT TRANSMEMBRANE METHOTREXATE PD003967: S11-E230, G327-S493, F262-W303	BLAST_PRODUM
18	4253888CD1	573	S8 S35 S77 S88 S106 S137 S229 S304 S321 S340 S379 T10 T27 T148 T194 T401 Y115	N78 N466	Transmembrane domains: G157-Q185 V189-R217 A242-G268 K269-M297 V314-S340 E343-K369 L377-D399 L407-S433 A469-L489 G493-D521 G531-H559 N-terminus is cytosolic	TMAP
					Divalent cation transporter domain: L199-S335, Y413-H559	HMMEF_PPFAM
19	7479974CD1	573	S69 S138 S198 S221 S261 S355 S473 S478 S509 S543 T24 T159 T176 T401 T406 T474 T499	N80 N174 N294 N491 N529 N558	Anion-transporting ATPase domain: L354-S573	HMMEF_PPFAM



Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
19					PLASMAID ARSENICAL PUMPDRIVING ATPASE HYDROLASE RESISTANCE ATPBINDING ARSA PD006335: D460-P568 NIFH/FRXC FAMILY DM00105 P08690 7-180: Y20-E191 DM00105 P08690 326-473: L333-K472 DM00105 P30632 17-190: K19-S167 ATP/GTP-binding site motif A (P-loop): G25-T32, G337-T344 Transmembrane domains: D8-S36, R44-I72, C96-R123, S133-I158, L171-F194 N-terminus is non-cytosolic POT family (proton/oligopeptide symporter) domain: G56-N141	BLAST_PRODOM  BLAST_DOMO  NOTIFS
20	7483850CD1	248	S42 S85 S121 S198 S215 S227 S233 S239		Transmembrane domains: D8-S36, R44-I72, C96-R123, S133-I158, L171-F194 N-terminus is non-cytosolic POT family (proton/oligopeptide symporter) domain: G56-N141	TMAP  HMMER_PPFAM
21	5508353CD1	761	S5 S62 S109 S115 S185 S312 S409 S476 S556 S706 S734 T28 T30 T78 T162 T201 T227 T335 T534 T674 T695 T738 Y189	N700 N732	Transmembrane domains: P84-K112, Y459-N487, P528-Y553, G564-K584, L592-C612, I624-Y652 N-terminus is cytosolic  E1-E2 ATPases phosphorylation site proteins: BL00154: D231-L271, T386-S409, K131-L141 P-type cation-transporting atpase superfamily signature PR00119: A247-D257 H+-transporting ATPase (proton pump) signature PR00120: T162-A180 ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATP-BINDING PROTEIN PROBABLE CALCIUM TRANSPORTING CALCIUM TRANSPORT PD004657: A423-P661 ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATP-BINDING PROBABLE PROTEIN CALCIUM TRANSPORTING CALCIUM TRANSPORT PD149930: C363-F422	BLIMPS_BLOCKS  BLIMPS_PRINTS  BLIMPS_PRINTS  BLAST_PRODOM  BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
21					ATPase; CALCIUM; TRANSPORTING; DM02405 P32660 318-1225: E45-N487 ATPase; CALCIUM; TRANSPORTING; DM02405 Q09891 206-1107: L56-N487 ATPase; CALCIUM; TRANSPORTING; DM02405 S51243 356-1267: E59-F486 ATPase; CALCIUM; TRANSPORTING; DM02405 P39524 236-1049: Y58-D287, L332-N487 ATP/GTP-binding site motif A (P-loop): G331-T338, A699-S706 Binding-protein-dependent transport system: A112-Y185 Transmembrane domains: A20-N48, V70-S98, T156-R183, Y185-A208 N-terminus is non-Cytosolic Binding-protein-dependent transport systems inner membrane component: V105-T161 PROTEIN TRANSPORT TRANSMEMBRANE PERMEASE MEMBRANE AMINO ACID INNER SYSTEM TRANSPORTER ABC PD001196: E7-G111 BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEMS INNER MEMBRANE DM00388 P10345 3-214: L9-R215 BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEMS INNER MEMBRANE DM00388 P42399 12-220: S12-R215 BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEMS INNER MEMBRANE DM00388 P45023 17-232: P14-R215 BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEMS INNER MEMBRANE DM00388 P42200 15-226: P14-R215 Binding-protein-dependent transport systems inner membrane component signature: Lil3-P141	BLAST_DOMO  BLAST_DOMO  BLAST_DOMO  BLAST_DOMO  MOTIFS
22	8543628CD1	219	S4 T148			HMMER_PPAM
23	7482754CD1	463	S249 S254 S413 S420 T451 T458		Transmembrane amino acid transporter protein: A48-C433	HMMER_PPAM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
23					Transmembrane domains: L24-N44, G54-V74, R93-R121, D148-R168, H180-T200, S249-S274, G298-R326, W353-M373, S377-P397 N-terminus is non-cytosolic Transmembrane four family signature PR00259: F59-V82, L73-A99, V302-V328 ACID AMINO PROTEIN TRANSPORTER PERMEASE TRANSMEMBRANE INTERGENIC REGION PUTATIVE PROLINE PD001875: F31-I320	TMAP  BLIMPS_PRINTS BLAST_PRODROM
24	3794818CD1	1043	S36 S153 S206 S372 S388 S402 S536 S633 S700 S709 S721 S747 S849 S867 S876 S891 S901 S1014 S1036 T157 T192 T210 T421 T491 T593 T604 T636 T641 T673 T696 T725 T916	N69 N344 N451 N465 N609 N786	Signal Peptide: M1-G21	HMMER
					signal_cleavage: M1-G22	SPSCAN
					Ligand-gated ion channel: H574-E852	HMMER_PFBAM
					Transmembrane domains: H574-R598, P640-V667, T823-L847 N terminus is non-cytosolic	TMAP
					NMDA receptor signature PR00177: F493-L521, T577-G602, L644-D671, F831-F855	BLIMPS_PRINTS
					R32184_2 IONOTROPIC GLUTAMATE RECEPTOR PD156309: A77-Y477	BLAST_PRODROM
					RECEPTOR GLUTAMATE SUBUNIT SIGNAL PRECURSOR CHANNEL IONIC TRANSMEMBRANE POSTSYNAPTIC PD000500: M570-E852	BLAST_PRODROM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypep- tide ID	Amino Acid Residues	Potential Phosphoryl- ation Sites	Potential Glycosyl- ation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
24					RECEPTOR SIGNAL GLUTAMATE SUBUNIT PROTEIN TRANSMEMBRANE CHANNEL IONIC PD000273: G478-A563, G728-V817 GLUTAMATE RECEPTOR DM00247 Q03391 640-919: T631-S901, C964-P980 DM00247 P35436 615-886: T631-P856 DM00247 Q01098 613-882: T631-E904 DM00393 Q01097 377-614: G387-F628	BLAST_PRODUM  BLAST_DOMO
25	4717525CD1	480	S4 S23 S56 S105 S176 S411 S418 T161 T170 T220 T302 T410 T469		Mitochondrial carrier proteins: M184-T276, G319-L413	HMMEP_PPFAM
					EF hand: Q117-H145, R13-L41, R81-L109	HMMEP_PPFAM
					Mitochondrial energy transfer proteins BL00215: V190-Q214, I369-G381	BLIMPS_BLOCKS
					Mitochondrial energy transfer proteins signature: I320-S371, K187-L241, V279-S331	PROFILES SCAN
					Mitochondrial carrier protein signature PR00926: Q188-T201, T201-V215, G244-E264, C333-Q351, Y379-Y397, G327-Q349	BLIMPS_PRINTS
					Graves disease carrier protein signature PR00928: Q274-Q294, P205-I225, Y263-S287, I369-V389	BLIMPS_PRINTS
					TRANSPORT TRANSMEMBRANE CARRIER INNER MITOCHONDRIAL ADP/ATP PD000117: K187-S466	BLAST_PRODUM
					MITOCHONDRIAL ENERGY TRANSFER PROTEINS DM00026 S57544 26-107: V190-I270 DM00026 Q01888 38-124: V190-I270	BLAST_DOMO
					EF-hand calcium-binding domain: D22-L34 D90-I102	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
26	509173CD1	518	S8 S38 S161 S240 S253 S331 S389 S510 T378	N6 N171 N371 N376	Transmembrane domains: G77-W97, K104-L124, N131-V151, A170-T191, Q206-E234, I264-P292, Q304-K332, E336-I364, A406-W434, P481-L509 N terminus is non-cytosolic.	TMAP
27	5945527CD1	501	S39 S66 S263 S267 S329 S421 T338 T418	N53 N62 N68	Transmembrane domains: V9-H37, E85-S105, L114-L134, L197-T225, G290-W317, V342-A362, Y379-G399, T433-L453, N460-L480 N terminus is cytosolic. glpT family of transport BL00942: T29-K41, N82-L124, W171-V190, F211-P247, E281-Y321, L339-D356	TMAP
					GLPT FAMILY OF TRANSPORTERS DM02439 P37948 1-403: K84-H244, L305-L446 DM02439 P09836 1-401: L87-E234, P295-A426, S16-K41 DM02439 P08194 1-403: L87-D256, Q251-A444, R22-I45	BLIMPS_BLOCKS
					DM02439 P37948 1-403: K84-H244, L305-L446 DM02439 P09836 1-401: L87-E234, P295-A426, S16-K41 DM02439 P08194 1-403: L87-D256, Q251-A444, R22-I45	BLAST_DOMO
28	6941124CD1	801	S21 S171 S305 S399 S442 S469 S564 T6 T20 T56 T130 T174 T198 T200 T400 T420 T438 T704 Y791	N52 N388 N455 N463	HCO3- transporter family: A571-L745, H240-F546, R119-L208	HMMER_PFAM
					Transmembrane domains: V282-G304, I330-I358, T373-M393, I403-I423, H478-Y499, V514-E542, S564-L592, T604-A630, S672-Y700, L748-I776 N terminus is non-cytosolic.	TMAP

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
28					Anion exchangers family BL00219: S448-T483, P243-V282, T289-D312, L342-F380, A382-Y429, Q481-S534, A567-W608, D609-E647, H653-F698, Y700-T743, L748-I787 Anion exchanger signature PR00165: G253-I275, K284-G304, A371-S390, A482-L501, I517-G537 ANION EXCHANGE TRANSMEMBRANE GLYCO-PROTEIN LIPOPROTEIN PALMITATE BICARBONATE COTRANSPORTER PD001455: T483-T743, F752-S457, R158-A237 BAND 3 ANION TRANSPORT PROTEIN DM02294 P48751 601-1229: A190-E798 DM02294 P02730 311-908: T483-E798 DM02294 A42497 403-1027: F252-E798 DM02294 P04920 602-1237: D248-E798	BLIMPS_BLOCKS
29	6972530CD1	344	S101 S133 S157 S319 S333 T117 T145 T196 T286 Y330	N223 N299	signal_cleavage: M1-A34 Signal Peptide: M9-A34 Transmembrane domains: R8-L36 V39-V63 Q71-T99 N terminus is non-cytosolic.	SPSCAN HMMER TMAP

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
30	6991750CD1	2701	S1063 S1231 S1270 S40 S1367 S1526 S551 S1603 S193 S1710 S1793 S223 S1824 S1866 S274 S1879 S230 S1893 S1915 S651 S1922 S307	N56 N218 N319 N330 N364 N396 N452 N480 N539 N607 N667 N691 N772 N828 N875 N885 N921 N949 N964 N988 N1135	Transmembrane domains: I965-S986, E1281-T1300, N1670-S1692, V1836-W1864, V2276-N2292, A2507-P2534 N terminus is non-cytosolic.	TMAP
			S1952 S1978 S461 S2029 S382 S2045 S2090 S2230 S495 S2388 S2438 S792 S2464 S790 S2488 S2496 S2506 S901 S2537 S809 S2545 S2573 S856 S2631	N1153 N1166 N1241 N1259 N1273 N1305 N1323 N1372 N1479 N1525 N1528 N1618 N1624 N1892 N1969 N2134 N2342 N2428 N2510 N2596 N2625	Adenosine and AMP deaminase signature . S2386-P2392	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
30			T122 T158 T276 T417 T581 T774 T881 T989 T1004 T1030 T1045 T1051 S941 T1058 T1090 S953 T1307 T1559 T1610 T1658 T1786 T1887 T1945 T2171 T2213 T2266 T2328 T2409 T2415 T2446 T2450 T2459 T2549 T2694 Y130		Sodium/Calcium Exchanger Chain DM05297 P48765 6-969; V263-I492 P=2.8-09	BLAST_DOMO
31	71726948CD1	610	S114 S269 S317 S375 S377 S563 S576 S602 T3 T41 T53 T74 T158 T312 T364 T483 T491 T557 Y554	N260 N481 N485 N606	Signal Peptide: M47-P73	HMME
					Sodium: solute symporter family: F45-G449	HMME_PPAM
					Transmembrane domains: P4-Y32, T53-G73, I84-L104, L121-Y148, G159-L179, V190-M210, R239-I256, R273-L301, S385-L405, A412-I432, S439-G459, N515-T543 N terminus is non-cytosolic.	TMAP
					Sodium: solute symporter BL00456: L22-S76, I98-V127, T158-G212	BLIMPS_BLOCKS
					Sodium: solute symporter family signatures: N155-A201	PROFILES SCAN



Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
31					TRANSMEMBRANE TRANSPORT PERMEASE SODIUM SYMPORT PROLINE GLYCOPROTEIN PD000991: F45-P234, L129-G449 SYMPORTER SODIUM IODIDE THYROID PD024705: A451-P552 SODIUM: SOLUTE SYMPORTER FAMILY DM00745 JC2382 3-485:Y15-W455 DM00745 P45174 3-495:T9-W455 DM00745 P31448 1-494:F18-G449 DM00745 P44963 1-483:V23-Y463	BLAST_PRODUM BLAST_PRODUM BLAST_DOMO
32	7487393CD1	552	S46 S60 S68 S143 S167 S276 S282 S408 S475 S537 T58 T133 T311 T323 T391 T526	N39 N56 N62 N102 N377	Sugar (and other) transporter: I18-V530  Transmembrane domains: V10-E38, K145-G164, I174-L202, M232-A252, Q262-S282, K345-I368, G375-L397, F412-L440, S475-L496, P497-I514 N terminus is non-cytosolic.	HMMEER_PPFAM TMAP
					ORGANIC TRANSPORTERLIKE TRANSPORT PROTEIN RENAL ANION TRANSPORTER CATIONIC KIDNEY SPECIFIC SOLUTE PD151320: N102-K145	BLAST_PRODUM

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
33/7484831CB1/2365	1-325, 94-240, 106-240, 106-242, 106-244, 106-245, 106-514, 106-515, 106-582, 106-586, 106-593, 106-597, 106-612, 109-245, 110-245, 128-770, 160-516, 164-267, 176-549, 181-440, 184-581, 234-317, 234-409, 252-770, 275-469, 275-473, 275-570, 275-636, 275-700, 336-1038, 338-1038, 377-543, 413-622, 414-1038, 495-1038, 504-764, 504-1009, 521-952, 521-1034, 521-1038, 523-680, 704-974, 781-1246, 824-1377, 884-1127, 904-1254, 916-1533, 940-1038, 958-1226, 978-1038, 990-1038, 1018-1637, 1031-1722, 1064-1463, 1070-1211, 1086-1149, 1110-1336, 1110-1375, 1110-1386, 1110-1387, 1110-1391, 1110-1462, 1113-1382, 1139-1622, 1144-1622, 1153-1629, 1178-1434, 1178-1749, 1210-1694, 1262-1309, 1445-1969, 1445-2000, 1455-1830, 1463-1630, 1478-1701, 1519-1790, 1526-1952, 1582-1838, 1669-1886, 1676-1846, 1772-2205, 1846-2117, 1847-2060, 1890-2209, 1934-2041, 1938-2315, 1950-2365
34/2477266CB1/3400	1-89, 1-306, 13-348, 13-665, 16-468, 23-270, 23-271, 23-404, 23-512, 27-306, 35-386, 35-589, 37-693, 124-795, 240-510, 317-524, 366-999, 459-996, 473-1144, 492-1041, 493-1060, 505-1019, 761-1360, 798-1025, 798-1268, 872-1360, 875-1143, 988-1278, 996-1252, 996-1612, 1024-1265, 1032-1623, 1054-1507, 1078-1341, 1192-1726, 1213-1608, 1218-1469, 1253-1798, 1332-1567, 1332-1800, 1345-1908, 1387-1693, 1419-1981, 1434-1737, 1436-1726, 1518-2165, 1524-2132, 1562-1836, 1775-2024, 1779-2022, 1797-2084, 1890-2052, 1890-2116, 1890-2334, 1890-2438, 1913-2196, 1942-2197, 1949-2528, 1952-2509, 1954-2029, 1981-2621, 1994-2104, 1994-2174, 1994-2252, 1994-2258, 1994-2259, 2039-2147, 2040-2319, 2062-2315, 2142-2732, 2205-2493, 2238-2536, 2265-2795, 2275-2544, 2277-2528, 2302-2718, 2337-2622, 2380-2920, 2444-2841, 2467-2711, 2472-2728, 2472-2737, 2472-2955, 2484-2753, 2492-2641, 2503-2746, 2512-3142, 2517-3095, 2555-3131, 2568-3246, 2590-2931, 2605-2903, 2643-2973, 2646-2931, 2656-2902, 2673-3284, 2725-3400, 2766-3390, 2766-3398, 2795-3082, 2800-3097, 2818-3039, 2818-3355, 2832-3051, 2832-3088, 2851-3007, 2851-3015, 2851-3028, 2854-3120, 2868-3336, 2871-3133, 2871-3195, 2884-3181, 2892-3104
35/3552033CB1/4458	1-1188, 405-1189, 422-1189, 447-1152, 479-1274, 507-1189, 511-1189, 515-1185, 523-1188, 535-1189, 540-1188, 547-1189, 550-1189, 560-1189, 561-1189, 570-1189, 571-1275, 579-1189, 585-1189, 586-1141, 602-1189, 602-1275, 629-1275, 631-1189, 635-1189, 636-1275, 638-1189, 640-1189, 643-1189, 653-1189, 654-1189, 655-1120, 675-1240, 695-1188, 699-1188, 700-1189, 701-1189, 708-1189, 735-1189, 743-1189, 746-1189, 748-1189, 749-1189, 752-1189, 761-1189, 762-1189, 763-1189, 768-1189, 775-1189, 780-1188, 820-1189, 844-1600, 887-1496, 970-1185, 1008-1841, 1023-1188, 1028-1189, 1028-1721, 1055-1574, 1108-1275, 1171-1275, 1491-2041, 1491-2086, 1551-2416, 1874-2534, 1874-2548, 1874-2561, 1874-2568, 1874-2574, 1874-2576, 1874-2580, 1875-2432, 1875-2436, 1875-2539, 1887-2564, 1909-2592, 1917-2507, 1940-2494, 1955-2542, 1959-2477, 1962-2428, 1964-2460, 1978-2424, 1988-2500, 1992-2500, 1994-2673, 1996-2500, 2014-2906, 2017-2437, 2025-2578, 2025-2580, 2033-2525, 2038-2492, 2045-2611, 2053-2644, 2066-2555, 2087-2691, 2104-2699, 2105-2514, 2147-2427,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
35	2171-2621, 2181-2699, 2196-2743, 2205-3023, 2207-2752, 2215-2959, 2221-3026, 2221-3058, 2222-3020, 2249-2765, 2290-3024, 2291-2959, 2295-2796, 2300-2958, 2320-3022, 2325-2763, 2384-3024, 2391-3024, 2406-3024, 2407-3024, 2414-3023, 2414-3024, 2420-3024, 2433-2718, 2433-2900, 2433-3024, 2433-3025, 2434-3024, 2435-3024, 2438-3024, 2441-3024, 2443-3024, 2446-3024, 2451-3023, 2455-3024, 2470-3024, 2479-3024, 2480-3024, 2482-3024, 2484-3024, 2488-3024, 2489-3024, 2491-3024, 2501-3024, 2508-3024, 2511-3024, 2522-3024, 2582-3024, 2583-3024, 2587-3024, 2590-3024, 2592-2877, 2592-3022, 2592-3024, 2596-3024, 2602-3279, 2612-3024, 2621-3063, 2621-3092, 2626-3279, 2632-3024, 2652-3024, 2653-3024, 2671-3015, 2671-3020, 2671-3022, 2671-3024, 2686-3024, 2688-3024, 2689-3024, 2718-3024, 2732-3279, 2734-3024, 2736-3279, 2763-3024, 2766-2830, 2770-3021, 2773-3203, 2802-3279, 2812-3004, 2831-3008, 2831-3037, 2831-3117, 2831-3220, 2831-3247, 2831-3275, 2831-3300, 2831-3490, 2842-3330, 2853-3278, 2866-3103, 2902-3270, 2998-3234, 2998-3456, 2998-3524, 2998-3550, 2998-3606, 3003-3715, 3032-3279, 3052-3811, 3078-3810, 3175-3935, 3182-3854,
36/4778139CB1/2722	3219-3792, 3224-3807, 3247-3811, 3267-3885, 3355-3811, 3359-3429, 3359-3811, 3389-4035, 3393-4035, 3402-3915, 3423-3982, 3427-4044, 3444-3811, 3463-4149, 3472-4150, 3485-4080, 3498-3527, 3510-3811, 3521-4030, 3547-3782, 3552-4035, 3552-4039, 3606-4267, 3606-4430, 3629-3871, 3629-3872, 3629-4047, 3631-4047, 3660-4047, 3673-4416, 3676-3848, 3703-3811, 3713-3864, 3714-3972, 3767-4458, 3792-4454, 3817-4084, 3820-4047, 3852-4458, 3857-4035, 3932-4144, 3934-4347, 3939-4152, 4201-4458, 4219-4458, 4295-4458
37/4787433CB1/1924	1-490, 1-545, 1-554, 1-567, 1-574, 1-575, 1-594, 1-606, 1-671, 1-1206, 5-2226, 157-394, 424-874, 433-874, 448-874, 459-874, 461-874, 467-874, 481-874, 498-874, 508-874, 563-874, 580-874, 1464-1728, 1464-1973, 1660-1896, 1660-1919, 1935-2155, 2166-2453, 2166-2722, 2178-2456
38/7483598CB1/1797	1-77, 1-100, 1-480, 154-1023, 513-772, 719-980, 719-1247, 729-984, 805-1016, 843-1529, 848-1456, 849-1459, 902-1530, 902-1552, 940-1543, 951-1593, 952-1426, 962-1646, 990-1548, 1001-1588, 1029-1577, 1047-1657, 1048-1442, 1059-1472, 1059-1569, 1064-1527, 1075-1646, 1076-1628, 1082-1734, 1091-1693, 1095-1701, 1099-1339, 1104-1681, 1104-1707, 1106-1674, 1112-1332, 1119-1583, 1123-1765, 1129-1683, 1169-1608, 1200-1626, 1207-1745, 1218-1832, 1230-1611, 1232-1669, 1244-1741, 1256-1758, 1265-1584, 1273-1758, 1280-1557, 1284-1765, 1303-1924, 1356-1592, 1356-1924, 1359-1585, 1359-1833, 1366-1917, 1382-1911, 1387-1913, 1518-1924, 1575-1909, 1776-1886
	1-808, 8-302, 13-417, 14-808, 179-808, 490-808, 491-808, 732-1017, 881-1226, 881-1227, 891-1105, 1157-1429, 1157-1700, 1184-1753, 1184-1760, 1197-1708, 1197-1754, 1197-1755, 1197-1757, 1197-1797, 1198-1757, 1199-1757, 1199-1797, 1200-1754, 1292-1707, 1295-1760, 1324-1759, 1359-1744, 1359-1749, 1359-1760, 1362-1760

Table 4

Polynucleotide SEQ ID NO./ Incye ID/ Sequence Length	Sequence Fragments
39/7484823CB1/3277	1-174, 64-1458, 1324-1458, 1326-1458, 1327-1458, 1340-1458, 1458-1660, 1459-1659, 1459-1660, 1459-2936, 1484-1651, 1506-1732, 1683-2339, 1700-2334, 1724-2174, 1724-2423, 1792-2400, 1804-2530, 1822-2278, 1834-2323, 1840-2516, 1878-2430, 1891-2244, 1922-2694, 1926-2592, 1931-2472, 1938-2431, 1940-2458, 1941-2472, 1944-2546, 1956-2566, 1962-2561, 1970-2530, 1976-2515, 1977-2497, 1980-2431, 2012-2660, 2020-2596, 2046-2671, 2049-2576, 2083-2578, 2095-2752, 2095-2763, 2098-2659, 2105-2794, 2114-2787, 2129-2908, 2162-2720, 2163-2689, 2225-3024, 2229-2763, 2241-2925, 2358-3057, 2374-3044, 2444-2930, 2455-3277, 2531-3023, 2557-3140, 2586-3078
40/143935CB1/2773	1-217, 1-462, 1-942, 16-601, 34-217, 39-566, 176-217, 206-667, 217-756, 217-769, 274-800, 288-919, 316-814, 316-956, 328-967, 334-597, 374-939, 377-855, 388-785, 388-930, 389-791, 407-678, 409-1068, 435-1053, 464-1123, 468-723, 475-1045, 524-1016, 527-684, 527-1039, 564-964, 564-1004, 564-1098, 565-896, 574-1137, 598-933, 660-1190, 660-1192, 663-1285, 690-967, 704-1368, 713-1134, 748-1007, 760-1016, 760-1032, 760-1210, 769-1455, 771-1097, 782-1415, 808-1391, 867-1133, 870-1407, 873-1460, 885-1501, 899-1078, 899-1158, 961-1613, 976-1240, 994-1459, 1006-1508, 1014-1386, 1017-1416, 1025-1452, 1044-1307, 1059-1249, 1059-1668, 1085-1705, 1098-1506, 1099-1346, 1099-1401, 1099-1545, 1149-1353, 1157-1592, 1157-1661, 1193-1716, 1202-1431, 1212-1753, 1250-1654, 1268-1570, 1285-1748, 1285-1841, 1285-1851, 1285-1919, 1285-1930, 1285-1945, 1285-1952, 1285-1963, 1285-1965, 1285-1979, 1292-1886, 1300-1453, 1305-1577, 1310-1858, 1316-1530, 1317-1569, 1317-1872, 1317-1915, 1353-1865, 1356-1488, 1372-1714, 1416-2049, 1421-1709, 1430-1847, 1437-1742, 1441-1991, 1500-1747, 1505-2163, 1515-1815, 1515-1911, 1538-2068, 1556-1827, 1560-1954, 1572-1793, 1573-1837, 1580-2192, 1585-2093, 1601-1839, 1601-1840, 1618-2132, 1632-1887, 1651-1754, 1686-1934, 1686-1950, 1686-1965, 1686-2298, 1688-1978, 1689-1955, 1734-2319, 1743-2254, 1743-2377, 1749-2032, 1755-1996, 1755-2346, 1762-2492, 1763-2071, 1774-2011, 1784-2152, 1795-2063, 1795-2066, 1797-2299, 1819-2416, 1826-2091, 1827-2517, 1829-2490, 1829-2508, 1839-2076, 1842-2348, 1868-2396, 1868-2421, 1873-2459, 1875-2415, 1877-2145, 1878-2563, 1880-2052, 1880-2415, 1880-2467, 1880-2470, 1897-2131, 1907-2180, 1908-2168, 1917-2209, 1919-2386, 1919-2470, 1922-2470, 1923-2179, 1936-2195, 1937-2180, 1941-2267, 1944-2627, 1945-2556, 1965-2538, 1980-2269, 1982-2504, 1983-2560, 2001-2303, 2005-2567, 2012-2470, 2021-2569, 2027-2289, 2027-2303, 2027-2487, 2039-2531, 2048-2277, 2058-2293, 2089-2569, 2095-2559, 2096-2559, 2104-2561, 2108-2580, 2110-2564, 2113-2559, 2115-2669, 2116-2564, 2117-2561, 2120-2422, 2122-2577, 2128-2439, 2129-2580, 2143-2559, 2145-2670, 2146-2561, 2151-2453, 2152-2559, 2153-2401, 2153-2572, 2154-2562, 2156-2580, 2168-2438, 2168-2550, 2169-2559, 2171-2563, 2188-2405, 2188-2559, 2196-2773, 2204-2472, 2220-2455, 2222-2559, 2230-2538, 2233-2471, 2233-2502, 2233-2773, 2235-2567, 2237-2572, 2243-2536, 2251-2564, 2279-2529, 2279-2543, 2296-2559, 2296-2773, 2299-2561, 2323-2630, 2332-2567, 2333-2621, 2346-2562

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
41/5923789CB1/2966	1-1985, 1337-1980, 1564-2177, 1960-2223, 1960-2460, 1960-2503, 1960-2622, 1973-2554, 2029-2390, 2087-2410, 2087-2686, 2087-2838, 2139-2966, 2161-2348, 2193-2420, 2205-2412, 2216-2420, 2227-2838, 2233-2838, 2271-2838, 2306-2833, 2314-2778, 2327-2838, 2337-2838, 2355-2767, 2355-2838, 2359-2838, 2377-2838, 2378-2838, 2396-2575
42/6046484CB1/4404	1-966, 355-1040, 456-1274, 477-1118, 520-1325, 722-1557, 817-1375, 859-1564, 940-1040, 1025-1441, 1165-1850, 1198-1314, 1314-1850, 1413-1850, 1417-4404, 1444-1850
43/7481427CB1/669	1-669
44/7483595CB1/1823	1-351, 309-664, 348-668, 541-1059, 817-889, 988-1355, 1060-1248, 1060-1663, 1277-1663, 1354-1411, 1354-1577, 1354-1636, 1355-1823
45/3788427CB1/2931	1-1173, 428-1018, 432-983, 463-733, 497-793, 649-1173, 707-1128, 771-1417, 842-1065, 975-1512, 1056-1310, 1056-1604, 1056-1746, 1056-1771, 1099-1681, 1114-1696, 1133-1562, 1197-1817, 1309-1798, 1343-1972, 1344-1884, 1363-1652, 1363-1912, 1364-1978, 1371-1910, 1392-2005, 1396-1803, 1410-1943, 1419-2005, 1498-2167, 1502-1723, 1537-2239, 1547-1822, 1548-1975, 1551-2230, 1558-2210, 1561-2168, 1569-2279, 1570-2232, 1591-2280, 1596-2135, 1602-1879, 1617-2134, 1620-2322, 1670-2289, 1706-2210, 1714-1926, 1755-2450, 1773-2391, 1786-2373, 1805-2270, 1821-2067, 1890-2464, 1968-2608, 1979-2567, 2127-2446, 2161-2439, 2230-2425, 2230-2677, 2230-2722, 2230-2752, 2230-2762, 2230-2821, 2245-2512, 2247-2378, 2247-2419, 2247-2838, 2275-2558, 2276-2822, 2314-2585, 2352-2630, 2394-2863, 2446-2930, 2466-2733, 2486-2734, 2490-2743, 2490-2770, 2578-2822, 2599-2834, 2603-2861, 2618-2891, 2628-2931
46/6972455CB1/1492	1-447, 1-576, 1-620, 1-622, 1-649, 255-890, 289-944, 496-1098, 670-1452, 772-1234, 856-1492, 876-1492, 907-1244, 933-1492, 1009-1492, 1061-1492
47/8077668CB1/2406	1-429, 108-429, 331-428, 331-652, 425-934, 430-768, 771-1250, 771-1254, 1212-1343, 1212-1573, 1321-1817, 1374-1664, 1750-2406, 1764-2198
48/55120485CB1/3686	1-63, 9-302, 9-397, 9-484, 9-565, 9-586, 9-628, 9-632, 9-635, 60-301, 60-304, 60-597, 64-238, 186-742, 226-730, 284-784, 335-779, 623-1272, 689-1272, 739-1272, 762-1272, 800-1204, 802-1273, 826-1272, 884-1272, 900-1272, 982-1915, 988-1272, 1024-1272, 1047-1272, 1104-1272, 1107-1272, 1160-1272, 1232-1272, 1262-1675, 1273-1653, 1273-1675, 1322-1675, 1323-1675, 1591-1675, 1634-1918, 1635-1918, 1674-2808, 1675-1918, 1675-2525, 2157-2256, 2526-2764, 2526-2928, 2526-2962, 2557-3036, 2644-2892, 2658-3194, 2664-2790, 2664-2909, 2672-2790, 2874-3078, 2888-3292, 2916-3215, 2916-3530, 2937-3614, 2950-3102, 3000-3665, 3007-3664, 3019-3093, 3090-3349, 3110-3673, 3399-3673, 3419-3681, 3476-3686

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
49/3112883CB1/2807	1-346, 20-455, 23-481, 32-509, 39-710, 83-240, 149-240, 170-281, 170-436, 170-440, 170-661, 170-697, 170-700, 170-725, 170-728, 170-729, 170-737, 170-784, 170-857, 175-705, 216-342, 235-394, 317-704, 408-1134, 428-569, 442-834, 458-1060, 472-912, 547-1270, 575-1134, 583-1254, 586-1204, 595-1258, 652-1234, 656-1231, 662-1226, 676-943, 676-1287, 698-1388, 701-968, 707-992, 707-1098, 762-1077, 762-1309, 765-1335, 788-1403, 798-1366, 805-1032, 816-1391, 825-1415, 843-1421, 857-1367, 866-1448, 930-1331, 936-1539, 946-1551, 947-1206, 949-1476, 1054-1469, 1074-1638, 1089-1638, 1129-1533, 1220-1571, 1232-1810, 1283-1529, 1325-1533, 1445-1674, 1481-2085, 1506-1764, 1515-2092, 1518-2142, 1530-1810, 1531-2141, 1552-2154, 1650-1908, 1650-2149, 1688-1961, 1731-2154, 1749-1984, 1808-2155, 1818-1960, 1838-2155, 2078-2807, 2108-2155, 2198-2426, 2296-2426, 2423-2575, 2423-2578, 2423-2583, 2423-2587, 2423-2625, 2423-2701, 2423-2742, 2423-2769, 2431-2561, 2444-2627, 2444-2769, 2484-2769, 2499-2769, 2502-2769, 2510-2769, 2516-2571, 2538-2769, 2546-2769, 2552-2769, 2578-2769, 2602-2769, 2611-2769, 2615-2769, 2625-2769
50/4253888CB1/2170	1-629, 105-746, 105-783, 106-600, 676-887, 676-1004, 715-1004, 825-936, 868-1104, 885-1482, 931-1612, 943-1612, 1112-1612, 1125-1348, 1125-1700, 1513-1988, 1532-1596, 1588-2038, 1596-2036, 1661-1893, 1813-2158, 1826-2058, 1826-2146, 1827-2073, 1827-2170, 1904-2162
51/7479974CB1/1722	1-1722, 251-1712
52/7483850CB1/1424	1-283, 1-533, 1-569, 1-578, 1-582, 1-583, 1-584, 5-584, 27-584, 59-285, 61-285, 64-285, 84-285, 139-283, 140-283, 284-347, 284-369, 284-403, 284-412, 284-431, 530-659, 594-1248, 594-1320, 594-1321, 594-1335, 594-1424, 599-1042
53/5508353CB1/3598	1-250, 12-553, 12-673, 12-704, 49-248, 49-568, 356-920, 356-929, 482-1038, 767-1173, 767-1201, 771-1247, 789-1080, 810-1477, 958-1217, 958-1486, 1010-1393, 1207-1393, 1348-1891, 1389-1776, 1399-2116, 1431-2008, 1434-2085, 1465-1961, 1682-2144, 1745-2115, 1783-1949, 1783-2318, 1792-1874, 1792-2229, 1833-2070, 1855-2127, 1855-2149, 1855-2185, 1855-2341, 1855-2432, 1958-2588, 1979-2104, 1979-2169, 2033-2712, 2064-2515, 2076-2331, 2177-2441, 2177-2708, 2238-2501, 2243-2883, 2245-2513, 2264-2537, 2301-2555, 2328-2636, 2363-2883, 2367-2906, 2381-2671, 2384-2626, 2384-2849, 2551-2763, 2551-2866, 2551-3054, 2713-2995, 2764-3247, 2799-3041, 2805-3598, 2819-3093, 2832-3264, 2841-3094
54/8543628CB1/1485	1-501, 1-502, 1-525, 1-563, 1-573, 1-580, 1-622, 1-751, 1-782, 1-812, 2-929, 11-37, 30-749, 33-854, 46-929, 65-788, 93-906, 144-883, 231-936, 261-1072, 491-1485, 509-1035
55/7482754CB1/1470	1-817, 73-267, 145-1266, 155-767, 631-734, 631-742, 641-1470, 713-766, 713-767, 793-1055, 908-1055, 908-1143, 1056-1143, 1056-1265, 1144-1265
56/3794818CB1/3132	1-2052, 1021-2199, 2071-3132, 2467-2702, 2555-2792, 2632-2702, 2722-2867

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
57/4717525CB1/1832	1-447, 69-292, 69-442, 69-460, 69-504, 69-506, 69-534, 69-625, 69-643, 69-675, 69-686, 69-693, 69-703, 69-705, 69-706, 69-732, 69-738, 80-434, 81-606, 83-464, 87-447, 147-1059, 157-413, 157-475, 181-443, 181-570, 187-802, 199-440, 210-537, 210-640, 237-535, 246-841, 255-517, 255-890, 300-830, 301-774, 327-990, 358-889, 377-1018, 379-671, 439-1025, 443-781, 445-997, 455-1059, 469-1060, 528-1013, 539-1059, 541-1048, 558-1059, 566-1003, 583-1059, 585-1059, 597-882, 599-1014, 619-864, 635-1059, 656-1059, 673-1059, 677-1059, 691-924, 692-1059, 711-1059, 717-1059, 720-1059, 728-1010, 735-1059, 741-1069, 760-1059, 765-1059, 779-1059, 794-1059, 796-1059, 805-1049, 814-1373, 821-941, 850-1059, 858-1069, 862-1059, 876-1059, 881-1059, 889-1059, 905-1049, 915-1059, 916-1059, 918-1059, 931-1059, 935-1059, 949-1046, 959-1059, 967-1059, 975-1059, 981-1049, 987-1059, 1002-1059, 1006-1059, 1008-1059, 1010-1382, 1015-1059, 1019-1059, 1047-1076, 1047-1093, 1047-1114, 1047-1139, 1047-1169, 1047-1173, 1047-1180, 1047-1184, 1047-1191, 1047-1198, 1047-1200, 1047-1206, 1047-1262, 1047-1269, 1047-1290, 1047-1305, 1047-1334, 1047-1335, 1047-1353, 1047-1355, 1047-1381, 1047-1382, 1047-1383, 1049-1382, 1053-1200, 1058-1382, 1063-1315, 1110-1381, 1128-1366, 1317-1832, 1478-1718, 1478-1811, 1482-1567
58/5091793CB1/1902	1-248, 1-544, 1-547, 1-567, 1-598, 1-648, 1-691, 1-720, 1-732, 1-737, 1-752, 1-796, 1-816, 1-833, 13-572, 48-884, 71-486, 71-550, 81-267, 82-549, 113-380, 113-549, 118-514, 127-558, 134-267, 159-986, 176-446, 265-528, 292-578, 375-663, 427-646, 593-879, 645-832, 645-884, 645-1125, 645-1149, 645-1164, 645-1233, 687-808, 769-1108, 806-1047, 806-1271, 806-1353, 807-974, 807-1318, 815-1109, 819-1004, 821-1461, 941-1108, 952-1641, 1041-1108, 1168-1561, 1213-1695, 1213-1742, 1215-1742, 1231-1742, 1233-1742, 1246-1878, 1250-1738, 1257-1614, 1315-1727, 1318-1847, 1349-1742, 1392-1902, 1411-1707, 1422-1742, 1426-1742, 1427-1742, 1438-1742, 1442-1742, 1458-1562, 1458-1628, 1458-1630, 1458-1633, 1458-1666, 1458-1680, 1458-1683, 1458-1693, 1458-1697, 1458-1742, 1476-1742, 1504-1902, 1511-1742, 1565-1742, 1576-1742
59/5945527CB1/2820	1-332, 1-444, 25-450, 31-538, 213-693, 347-617, 451-749, 451-1061, 785-1155, 792-1095, 895-942, 942-1056, 942-1096, 1033-1104, 1033-1199, 1044-1315, 1181-1275, 1181-1423, 1256-1526, 1280-1477, 1280-1905, 1389-1625, 1546-1821, 1546-2082, 1547-1589, 1547-1590, 1547-1611, 1547-1644, 1547-1669, 1547-1680, 1547-1681, 1547-1698, 1547-1717, 1547-1720, 1547-1724, 1547-1734, 1547-1757, 1547-1775, 1547-1776, 1547-1844, 1547-1867, 1547-1912, 1547-1957, 1547-2007, 1547-2042, 1547-2055, 1547-2071, 1547-2269, 1560-1840, 1627-2238, 1648-2068, 1656-1920, 1661-2365, 1667-2189, 1675-2354, 1698-2330, 1761-2039, 1776-2056, 1778-2037, 1778-2238, 1793-2339, 1814-2063, 1822-2494, 1837-2466, 1856-2516, 1868-2572, 1881-2465, 1990-2655, 1995-2567, 2050-2498, 2069-2668, 2134-2660, 2187-2507, 2187-2820, 2220-2820, 2222-2497, 2226-2492, 2240-2668, 2264-2668, 2271-2536, 2316-2668, 2384-2820, 2395-2654, 2395-2676, 2398-2668, 2401-2556, 2410-2558, 2745-2811, 2747-2820

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
60/6941124CB1/3920	1-552, 54-555, 384-722, 455-729, 548-1246, 726-938, 728-1114, 1033-1856, 1203-1438, 1203-1630, 1203-1701, 1260-1857, 1374-1642, 1414-1642, 1502-1642, 1615-2327, 1642-1992, 1642-2002, 1672-2155, 1672-2262, 1673-2328, 1673-2342, 1875-2328, 1932-2328, 1976-2483, 2070-2653, 2086-2328, 2169-2324, 2211-2328, 2296-2323, 2378-2625, 2378-2749, 2378-2816, 2378-2853, 2378-2866, 2378-2879, 2378-2882, 2378-2887, 2378-2908, 2378-2957, 2378-2964, 2378-2970, 2378-2971, 2378-2986, 2378-3019, 2378-3084, 2378-3087, 2378-3095, 2378-3136, 2378-3168, 2379-2852, 2400-3107, 2403-2520, 2403-2647, 2457-3035, 2474-3042, 2494-3206, 2512-3231, 2519-3086, 2521-2983, 2600-3107, 2600-3348, 2612-3344, 2631-3070, 2635-3311, 2663-3259, 2663-3415, 2758-3371, 2786-3425, 2787-2932, 2787-3205, 2793-3341, 2796-3360, 2824-3517, 2833-3458, 2852-3443, 2853-3373, 2856-3465, 2870-3494, 2881-3570, 2888-3404, 2888-3466, 2900-3526, 2907-3255, 2911-3421, 2911-3470, 2911-3479, 2912-3614, 2929-3314, 2938-3443, 2942-3443, 2954-3472, 2967-3537, 2989-3602, 3001-3247, 3004-3717, 3006-3704, 3014-3090, 3049-3536, 3058-3795, 3073-3813, 3078-3670, 3078-3807, 3094-3833, 3096-3785, 3100-3833, 3119-3833, 3168-3417, 3168-3575, 3168-3596, 3168-3838, 3168-3892, 3168-3919, 3169-3443, 3195-3833, 3220-3724, 3229-3856, 3230-3804, 3251-3549, 3254-3506, 3257-3801, 3260-3814, 3288-3832, 3289-3592, 3316-3833, 3333-3832, 3340-3833, 3348-3833, 3363-3840, 3391-3897, 3400-3833, 3402-3653, 3402-3832, 3409-3839, 3427-3830, 3427-3833, 3449-3720, 3464-3920, 3468-3783, 3477-3761, 3479-3727, 3479-3885, 3479-3920, 3500-3747, 3508-3733, 3521-3760, 3523-3910, 3530-3871, 3550-3690, 3581-3920, 3678-3910, 3701-3918, 3720-3906, 3746-3920, 3747-3894, 3747-3920, 3763-3910, 3786-3920, 3828-3920
61/6972530CB1/1333	1-549, 1-566, 391-966, 421-966, 711-1333
62/6991750CB1/8487	1-724, 32-456, 52-627, 78-723, 105-694, 135-592, 135-637, 143-723, 172-683, 175-420, 175-469, 175-496, 175-500, 175-541, 175-684, 175-764, 175-767, 175-784, 175-918, 187-677, 222-918, 254-723, 258-918, 305-918, 310-720, 358-918, 370-723, 386-918, 432-674, 432-918, 487-918, 847-1242, 847-1290, 895-1285, 896-3353, 917-1415, 917-1416, 917-1417, 920-1414, 1232-1739, 1232-1740, 1233-1739, 1236-1739, 1254-1898, 1266-1739, 1290-1834, 1290-1861, 1290-1868, 1581-2120, 1581-2121, 1584-2121, 1587-2121, 1950-2121, 2085-2762, 2088-2762, 2089-2762, 2090-2744, 2090-2754, 2090-2756, 2090-2758, 2090-2762, 2094-2762, 2152-2938, 2248-2900, 2269-2797, 2269-2949, 2993-3677, 3458-4151, 3585-8102, 3887-4151, 3931-4242, 4024-4540, 4024-4650, 4024-4651, 4027-4643, 4027-4648, 4027-4650, 4027-4651, 4903-5354, 4903-5386, 4903-5483, 4903-5507, 4903-5517, 4903-5525, 4903-5534, 4903-5548, 4903-5553, 4903-5561, 4903-5565, 4903-5567, 4903-5574, 5176-5434, 5340-6117, 5356-6117, 5364-6113, 5368-6117, 5393-6117, 5404-6117, 5409-6117, 5410-6117, 5417-6117, 5432-6117, 5434-6117, 5439-6117, 5449-6117, 6257-6768, 6398-6543, 6961-7140, 7135-7747, 7140-7955,



Table 4

Polynucleotide SEQ ID NO./ Incye ID/ Sequence Length	Sequence Fragments
62	7508-7900, 7614-8041, 7615-8102, 7843-8102, 7843-8209, 7843-8227, 7843-8236, 7843-8241, 7843-8262, 7843-8280, 7843-8341, 7843-8389, 7843-8430, 7843-8455, 7843-8465, 7843-8485, 7843-8486, 7843-8487
63/71726948CB1/3264	1-486, 4-441, 23-504, 24-504, 253-457, 286-504, 339-860, 355-823, 502-743, 502-1025, 550-856, 574-1259, 784-1237, 1104-1359, 1104-1362, 1104-1639, 1104-1654, 1104-1812, 1116-1681, 1360-1860, 1365-1834, 1366-1834, 1369-1881, 1419-2213, 1424-1890, 1523-2245, 1562-2184, 1724-2515, 1819-2533, 1838-2430, 1888-2228, 1888-2456, 1906-2469, 1931-2199, 1933-2190, 1991-2644, 1993-2507, 2008-2693, 2036-2682, 2046-2496, 2057-2684, 2172-2809, 2173-2809, 2268-2813, 2271-2813, 2301-2809, 2312-3014, 2328-2953, 2351-2751, 2397-3204, 2397-3261, 2404-2835, 2406-2835, 2444-3013, 2449-2716, 2466-3075, 2492-2960, 2536-3140, 2549-3201, 2574-3173, 2582-2743, 2609-2975, 2762-3237, 2807-3250, 2813-3244, 2852-3060, 3000-3249, 3113-3256, 3113-3264, 3114-3261
64/7487393CB1/1659	1-402, 1-1659, 307-506, 416-506, 547-860, 547-863, 547-867, 1147-1598

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
33	7484831CB1	LIVRNON08
34	2477266CB1	LIVRNON08
35	3552033CB1	BRAHTDK01
36	4778139CB1	PROSTUS19
37	4787433CB1	PGANNOT01
38	7483598CB1	BRAITUT29
39	7484823CB1	TESTNOC01
40	143935CB1	BRACDIK08
41	5923789CB1	BRAIFET02
42	6046484CB1	BRACNOK02
44	7483595CB1	TESTNOC01
45	3788427CB1	BONEUNR01
46	6972455CB1	BMARUNR02
47	8077668CB1	ADRETUE02
48	55120485CB1	BRAITUT29
49	3112883CB1	BRSTNOT03
50	4253888CB1	ADRETUE02
52	7483850CB1	LIVRDIT06
53	5508353CB1	NERDITDN03
54	8543628CB1	BMARUNR02
55	7482754CB1	PTHYTMN05
56	3794818CB1	KIDEUNE02
57	4717525CB1	KIDEUNE02
58	5091793CB1	LUNGTUT08
59	5945527CB1	SINTNOR01
60	6941124CB1	FTUBTUR01
61	6972530CB1	BMARUNR02
62	6991750CB1	BRAIFER06
63	71726948CB1	KIDNNOT32

Table 6

Library	Vector	Library Description
ADRETUB02	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from right adrenal tumor tissue removed from a 49-year-old Caucasian male during unilateral adrenalectomy. Pathology indicated adrenal cortical carcinoma comprising nearly the entire specimen. The tumor was attached to the adrenal gland which showed mild cortical atrophy. The tumor was encapsulated, being surrounded by a thin (1-3 mm) rim of connective tissue. The patient presented with adrenal cancer, abdominal pain, pyrexia of unknown origin, and deficiency anemia. Patient history included benign hypertension. Previous surgeries included adenotomylectomy. Patient medications included aspirin, calcium, and iron. Family history included atherosclerotic coronary artery disease in the mother; cerebrovascular accident and atherosclerotic coronary artery disease in the father; and benign hypertension in the grandparent(s).
BMARUNR02	PIGEN	This random primed library was constructed using RNA isolated from an untreated SH-SY5Y cell line derived from bone marrow neuroblastoma tumor cells removed from a 4-year-old Caucasian female.
BONEUNR01	PCDNA2.1	This random primed library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from an untreated MG-63 cell line derived from an osteosarcoma tumor removed from a 14-year-old Caucasian male (donor A) and using mRNA isolated from sacral bone tumor tissue removed from an 18-year-old Caucasian female (donor B) during an exploratory laparotomy and soft tissue excision. Pathology indicated giant cell tumor of the sacrum in donor B. Donor B's history included pelvic joint pain, constipation, urinary incontinence, unspecified abdominal/pelvic symptoms, and a pelvic soft tissue malignant neoplasm. Family history included prostate cancer in donor B.
BRACDIK08	PSPORT1	This amplified and normalized library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 57-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day for 40 years).

Table 6

Library	Vector	Library Description
BRAACNOK02	PSPORT1	This amplified and normalized library was constructed using RNA isolated from posterior cingulate tissue removed from an 85-year-old Caucasian female who died from myocardial infarction and retroperitoneal hemorrhage. Pathology indicated atherosclerosis, moderate to severe, involving the circle of Willis, middle cerebral, basilar and vertebral arteries; infarction, remote, left dentate nucleus; and amyloid plaque deposition consistent with age. There was mild to moderate leptomeningeal fibrosis, especially over the convexity of the frontal lobe. There was mild generalized atrophy involving all lobes. The white matter was mildly thinned. Cortical thickness in the temporal lobes, both maximal and minimal, was slightly reduced. The substantia nigra pars compacta appeared mildly depigmented. Patient history included COPD, hypertension, and recurrent deep venous thrombosis. 6.4 million independent clones from this amplified library were normalized in one round using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791.
BRAHTDK01	PSPORT1	This amplified and normalized library was constructed using pooled RNA isolated from archaocortex, anterior and posterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver. 7.6x10 <sup>5</sup> independent clones from this amplified library were normalized in 1 round using conditions adapted Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAIFER06	PCDNA2.1	This random primed library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation. Serologies were negative.
BRAIFET02	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
BRAITUT29	pINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the parietal lobe of a 43-year-old female during excision of a cerebral meningeal lesion. Pathology indicated high grade glioma. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, benign hypertension, and hyperlipidemia.

Table 6

Library	Vector	Library Description
BRSTNOT03	PSPORT1	Library was constructed using RNA isolated from diseased breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma. Patient history included kidney infection and condyloma acuminatum. Family history included benign hypertension, hyperlipidemia and a malignant neoplasm of the colon.
FTUBTUR01	PCDNA2.1	This random primed library was constructed using RNA isolated from fallopian tube tumor tissue removed from an 85-year-old Caucasian female during bilateral salpingo-oophorectomy and hysterectomy. Pathology indicated poorly differentiated mixed endometrioid (80%) and serous (20%) adenocarcinoma, which was confined to the mucosa without mural involvement. Endometrioid carcinoma in situ was also present. Pathology for the associated uterus tumor indicated focal endometrioid adenocarcinoma in situ and moderately differentiated invasive adenocarcinoma arising in an endometrial polyp. Metastatic endometrioid and serous adenocarcinoma was present at the cul-de-sac tumor. Patient history included medullary carcinoma of the thyroid and myocardial infarction.
KIDEUNE02	pINCY	This 5' biased random primed library was constructed using RNA isolated from an untreated transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue (Invitrogen). The cells were transformed with adenovirus 5 DNA.
KIDNNOT32	pINCY	Library was constructed using RNA isolated from kidney tissue removed from a 49-year-old Caucasian male who died from an intracranial hemorrhage and cerebrovascular accident. Patient history included tobacco abuse.
LIVRDIT06	pINCY	This library was constructed using RNA isolated from diseased liver tissue removed from a 35-year-old Caucasian male during needle biopsy of the liver. Patient history included hepatitis C.
LIVRNON08	pINCY	This normalized library was constructed from 5.7 million independent clones from a pooled liver tissue library. Starting RNA was made from pooled liver tissue removed from a 4-year-old Hispanic male who died from anoxia and a 16 week female fetus who died after 16-weeks gestation from anencephaly. Serologies were positive for cytomegalovirus in the 4-year-old. Patient history included asthma in the 4-year-old. Family history included taking daily prenatal vitamins and mitral valve prolapse in the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6

Library	Vector	Library Description
LUNGTUT08	pNCY	Library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fiberoptic bronchoscopy. Pathology indicated a grade 3 adenocarcinoma. Patient history included atherosclerotic coronary artery disease, an acute myocardial infarction, rectal cancer, an asymptomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, and lung cancer, type II diabetes, atherosclerotic coronary artery disease, and an acute myocardial infarction.
NERDTDN03	pNCY	This normalized dorsal root ganglion tissue library was constructed from 1.05 million independent clones from a dorsal root ganglion tissue library. Starting RNA was made from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema, acute bronchopneumonia, bilateral pleural effusions, pericardial effusion, and malignant lymphoma (natural killer cell type). The patient presented with pyrexia of unknown origin, malaise, fatigue, and gastrointestinal bleeding. Patient history included probable cytomegalovirus infection, liver congestion, and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia of the left lung, natural killer cell lymphoma of the pharynx, Bell's palsy, and tobacco and alcohol abuse. Previous surgeries included colonoscopy, closed colon biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy. Patient medications included Diflucan (fluconazole), Deltasone (prednisone), hydrocodone, Lortab, Alprazolam,
		Reazodone, ProMace-Cytabom, Etoposide, Cisplatin, Cytarabine, and dexamethasone. The patient received radiation therapy and multiple blood transfusions. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
PGANNOT01	PSPORT1	Library was constructed using RNA isolated from paraganglionic tumor tissue removed from the intra-abdominal region of a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and was associated with a grade 2 renal cell carcinoma, clear cell type, which did not penetrate the capsule. Surgical margins were negative for tumor.

Table 6

Library	Vector	Library Description
PROSTUS19	pINCY	This subtracted prostate tumor tissue library was constructed using 2.36 million clones from the PROSTUT13 library and was subjected to two round sof subtraction hybridization with 2.36 million clones from EPIPNOT01 library. The starting library for subtraction was constructed using RNA isolated from prostate tumor tissue removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenocarcinoma (Gleason grade 3+3) involving the prostate peripherally with invasion of the capsule. Adenofibromatous hyperplasia was present. The patient presented with elevated prostate-specific antigen (PSA). Patient history included diverticulitis of colon, asbestosis, and thrombophlebitis. Family history included benign hypertension, multiple myeloma, hyperlipidemia, and rheumatoid arthritis. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954 and Bonaldo, et al. Genome Research (1996) 6:791.
PTHYTMN05	pINCY	Library was constructed using RNA isolated from parathyroid tissue removed from a 44-year-old Caucasian male during a partial parathyroidectomy. Pathology for the matched tumor tissue indicated parathyroid carcinoma (grade 1 of 4) forming a partially cystic tan mass. Both capsular and vascular invasion were present. The patient presented with unspecified parathyroid disorder and calcium metabolism disorder. Patient history included kidney calculus and obesity. Previous surgeries included vasectomy and parathyroid surgery. Family history included emphysema in the mother; type II diabetes in the father; and type I diabetes and hyperlipidemia in the sibling(s).
SINTNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.
TESTNOC01	PBLUESCRIPT	This large size fractionated library was constructed using RNA isolated from testicular tissue removed from a pool of eleven, 10 to 61-year-old Caucasian males.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta B value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater, fastx B value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater



Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielsen, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Somhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
  - a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32,
  - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32,
  - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and
  - d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
  - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 1-32.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64,
- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64,
- c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain

reaction amplification, and

- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

5 17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

10

19. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition of claim 17.

15

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and  
b) detecting agonist activity in the sample.

20

21. A composition comprising an agonist compound identified by a method of claim 20 and  
a pharmaceutically acceptable excipient.

a

22. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment a  
25 composition of claim 21.

25

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide  
of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and  
30 b) detecting antagonist activity in the sample.

30

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

35

25. A method for treating a disease or condition associated with overexpression of

functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of

5 claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

10

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- 15 b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the
- 20 activity of the polypeptide of claim 1.

20

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

25

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying
- 30 amounts of the compound and in the absence of the compound.

30

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe
- 35 comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under

35

conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

- c) quantifying the amount of hybridization complex, and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A diagnostic test for a condition or disease associated with the expression of TRICH in biological sample, the method comprising:

- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a F(ab')<sub>2</sub> fragment, or
- e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the

composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- 5 a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- 10 c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

37. A polyclonal antibody produced by a method of claim 36.

15 38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- 20 a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- 25 d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

30 40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

35 42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32 in a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32 from a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.



49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

5 50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

10 51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

15 53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

20 55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

25 56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

30 58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

35 60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

5 63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

10

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

15

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

20

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

25

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

30

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

35

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.
80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
- 5 81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.
82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.
83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.
- 10 84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.
85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.
- 15 86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.
87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.
88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
20 NO:33.
89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:34.
- 25 90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:35.
91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:36.
- 30 92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:37.
93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
35 NO:38.

94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:39.

5 95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:40.

96. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:41.

10 97. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:42.

98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:43.

15 99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:44.

20 100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:45.

101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:46.

25 102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:47.

103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:48.

30 104. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:49.

35 105. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:50.

106. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:51.
- 5 107. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:52.
108. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:53.
- 10 109. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:54.
110. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:55.
- 15 111. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:56.
112. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:57.
- 20 113. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:58.
- 25 114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:59.
115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:60.
- 30 116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:61.
- 35 117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:63.

5 119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:64.

<110> INCYTE GENOMICS, INC.  
 LEE, Ernestine A.  
 BAUGHN, Mariah R.  
 YUE, Henry  
 DING, Li  
 RAUMANN, Brigitte E.  
 HAFALIA, April J.A.  
 KHAN, Farrah A.  
 NGUYEN, Danniel B.  
 ELLIOTT, Vicki S.  
 RAMKUMAR, Jayalaxmi  
 WALIA, Narinder K.  
 ISON, Craig H.  
 LU, Yan  
 GANDHI, Ameena R.  
 WARREN, Bridget A.  
 DUGGAN, Brendan M.  
 TRIBOULEY, Catherine M.  
 BURFORD, Neil  
 LU, Dyung Aina M.  
 LAL, Preeti  
 YAO, Monique G.  
 XU, Yuming  
 BRUNS, Christopher M.  
 THANGAVELU, Kavitha  
 SWARNAKAR, Anita  
 TANG, Y. Tom  
 AZIMZAI, Yalda  
 THORNTON, Michael  
 ARVIZU, Chandra  
 POLICKY, Jennifer L.

<120> TRANSPORTERS AND ION CHANNELS

<130> PI-0314 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/254,303; 60/256,190; 60/257,504; 60/261,546; 60/262,832;  
 60/264,377; 60/266,019

<151> 2000-12-08; 2000-12-15; 2000-12-21; 2001-01-12; 2001-01-19;  
 2001-01-26; 2000-02-02

<160> 64

<170> PERL Program

<210> 1

<211> 726

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7484831CD1

<400> 1

Met	Glu	Pro	Gly	Asp	Ala	Ala	Arg	Pro	Gly	Ser	Gly	Arg	Ala	Thr
1				5					10					15
Gly	Ala	Pro	Pro	Pro	Arg	Leu	Leu	Leu	Leu	Pro	Leu	Leu	Leu	Gly
				20					25					30
Trp	Gly	Leu	Arg	Val	Ala	Ala	Ala	Ala	Ser	Ala	Ser	Ser	Ser	Gly

				35					40					45
Ala	Ala	Ala	Glu	Asp	Ser	Ser	Ala	Met	Glu	Glu	Leu	Ala	Thr	Glu
				50					55					60
Lys	Glu	Ala	Glu	Glu	Ser	His	Arg	Gln	Asp	Ser	Val	Ser	Leu	Leu
				65					70					75
Thr	Phe	Ile	Leu	Leu	Leu	Thr	Leu	Thr	Ile	Leu	Thr	Ile	Trp	Leu
				80					85					90
Phe	Lys	His	Arg	Arg	Val	Arg	Phe	Leu	His	Glu	Thr	Gly	Leu	Ala
				95					100					105
Met	Ile	Tyr	Gly	Leu	Ile	Val	Gly	Val	Ile	Leu	Arg	Tyr	Gly	Thr
				110					115					120
Pro	Ala	Thr	Ser	Gly	Arg	Asp	Lys	Ser	Leu	Ser	Cys	Thr	Gln	Glu
				125					130					135
Asp	Arg	Ala	Phe	Ser	Thr	Leu	Leu	Val	Asn	Val	Ser	Gly	Lys	Phe
				140					145					150
Phe	Glu	Tyr	Thr	Leu	Lys	Gly	Glu	Ile	Ser	Pro	Gly	Lys	Ile	Asn
				155					160					165
Ser	Val	Glu	Gln	Asn	Asp	Met	Leu	Arg	Lys	Val	Thr	Phe	Asp	Pro
				170					175					180
Glu	Val	Phe	Phe	Asn	Ile	Leu	Leu	Pro	Pro	Ile	Ile	Phe	His	Ala
				185					190					195
Gly	Tyr	Ser	Leu	Lys	Lys	Arg	His	Phe	Phe	Arg	Asn	Leu	Gly	Ser
				200					205					210
Ile	Leu	Ala	Tyr	Ala	Phe	Leu	Gly	Thr	Ala	Val	Ser	Cys	Phe	Ile
				215					220					225
Ile	Gly	Asn	Leu	Met	Tyr	Gly	Val	Val	Lys	Leu	Met	Lys	Ile	Met
				230					235					240
Gly	Gln	Leu	Ser	Asp	Lys	Phe	Tyr	Tyr	Thr	Asp	Cys	Leu	Phe	Phe
				245					250					255
Gly	Ala	Ile	Ile	Ser	Ala	Thr	Asp	Pro	Val	Thr	Val	Leu	Ala	Ile
				260					265					270
Phe	Asn	Glu	Leu	His	Ala	Asp	Val	Asp	Leu	Tyr	Ala	Leu	Leu	Phe
				275					280					285
Gly	Glu	Ser	Val	Leu	Asn	Asp	Ala	Val	Ala	Ile	Val	Leu	Ser	Ser
				290					295					300
Ser	Ile	Val	Ala	Tyr	Gln	Pro	Ala	Gly	Leu	Asn	Thr	His	Ala	Phe
				305					310					315
Asp	Ala	Ala	Ala	Phe	Phe	Lys	Ser	Val	Gly	Ile	Phe	Leu	Gly	Ile
				320					325					330
Phe	Ser	Gly	Ser	Phe	Thr	Met	Gly	Ala	Val	Thr	Gly	Val	Val	Thr
				335					340					345
Ala	Leu	Val	Thr	Lys	Phe	Thr	Lys	Leu	His	Cys	Phe	Pro	Leu	Leu
				350					355					360
Glu	Thr	Ala	Leu	Phe	Phe	Leu	Met	Ser	Trp	Ser	Thr	Phe	Leu	Leu
				365					370					375
Ala	Glu	Ala	Cys	Gly	Phe	Thr	Gly	Val	Val	Ala	Val	Leu	Phe	Cys
				380					385					390
Gly	Ile	Thr	Gln	Ala	His	Tyr	Thr	Tyr	Asn	Asn	Leu	Ser	Val	Glu
				395					400					405
Ser	Arg	Ser	Arg	Thr	Lys	Gln	Leu	Phe	Glu	Val	Leu	His	Phe	Leu
				410					415					420
Ala	Glu	Asn	Phe	Ile	Phe	Ser	Tyr	Met	Gly	Leu	Ala	Leu	Phe	Thr
				425					430					435
Phe	Gln	Lys	His	Val	Phe	Ser	Pro	Ile	Phe	Ile	Ile	Gly	Ala	Phe
				440					445					450
Val	Ala	Ile	Phe	Leu	Gly	Arg	Ala	Ala	His	Ile	Tyr	Pro	Leu	Ser
				455					460					465
Phe	Phe	Leu	Asn	Leu	Gly	Arg	Arg	His	Lys	Ile	Gly	Trp	Asn	Phe
				470					475					480
Gln	His	Met	Met	Met	Phe	Ser	Gly	Leu	Arg	Gly	Ala	Met	Ala	Phe
				485					490					495
Ala	Leu	Ala	Ile	Arg	Asp	Thr	Ala	Ser	Tyr	Ala	Arg	Gln	Met	Met
				500					505					510



Phe	Thr	Thr	Thr	Leu	Ile	Val	Phe	Phe	Thr	Val	Trp	Ile	Ile
				515				520					525
Gly	Gly	Gly	Thr	Thr	Pro	Met	Leu	Ser	Trp	Leu	Asn	Ile	Arg
				530					535				540
Gly	Val	Glu	Glu	Pro	Ser	Glu	Glu	Asp	Gln	Asn	Glu	His	His
				545					550				555
Gln	Tyr	Phe	Arg	Val	Gly	Val	Asp	Pro	Asp	Gln	Asp	Pro	Pro
				560					565				570
Asn	Asn	Asp	Ser	Phe	Gln	Val	Leu	Gln	Gly	Asp	Gly	Pro	Asp
				575					580				585
Ala	Arg	Gly	Asn	Arg	Thr	Lys	Gln	Glu	Ser	Thr	Trp	Ile	Phe
				590					595				600
Leu	Trp	Tyr	Ser	Phe	Asp	His	Asn	Tyr	Leu	Lys	Pro	Ile	Leu
				605					610				615
His	Ser	Gly	Pro	Pro	Leu	Thr	Thr	Thr	Leu	Pro	Ala	Trp	Cys
				620					625				630
Leu	Leu	Ala	Arg	Cys	Leu	Thr	Ser	Pro	Gln	Val	Tyr	Asp	Asn
				635					640				645
Glu	Pro	Leu	Arg	Glu	Glu	Asp	Ser	Asp	Phe	Ile	Leu	Thr	Glu
				650					655				660
Asp	Leu	Thr	Leu	Thr	Tyr	Gly	Asp	Ser	Thr	Val	Thr	Ala	Asn
				665					670				675
Ser	Ser	Ser	Ser	His	Thr	Ala	Ser	Thr	Ser	Leu	Glu	Gly	Ser
				680					685				690
Arg	Thr	Lys	Ser	Ser	Ser	Glu	Glu	Val	Leu	Glu	Arg	Asp	Leu
				695					700				705
Met	Gly	Asp	Gln	Lys	Val	Ser	Ser	Arg	Gly	Thr	Arg	Leu	Val
				710					715				720
Pro	Leu	Glu	Asp	Asn	Ala								
				725									

<210> 2

<211> 1081

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2477266CD1

<400> 2

Met	Glu	Ser	Ala	Gly	Leu	Glu	Gln	Leu	Leu	Arg	Glu	Leu	Leu	Leu
1				5					10					15
Pro	Asp	Thr	Glu	Arg	Ile	Arg	Arg	Ala	Thr	Glu	Gln	Leu	Gln	Ile
				20					25					30
Val	Leu	Arg	Ala	Pro	Ala	Ala	Leu	Pro	Ala	Leu	Cys	Asp	Leu	Leu
				35					40					45
Ala	Ser	Ala	Ala	Asp	Pro	Gln	Ile	Arg	Gln	Phe	Ala	Ala	Val	Leu
				50					55					60
Thr	Arg	Arg	Arg	Leu	Asn	Thr	Arg	Trp	Arg	Arg	Leu	Ala	Ala	Glu
				65					70					75
Gln	Arg	Glu	Ser	Leu	Lys	Ser	Leu	Ile	Leu	Thr	Ala	Leu	Gln	Arg
				80					85					90
Glu	Thr	Glu	His	Cys	Val	Ser	Leu	Ser	Leu	Ala	Gln	Leu	Ser	Ala
				95					100					105
Thr	Ile	Phe	Arg	Lys	Glu	Gly	Leu	Glu	Ala	Trp	Pro	Gln	Leu	Leu
				110					115					120
Gln	Leu	Leu	Gln	His	Ser	Thr	His	Ser	Pro	His	Ser	Pro	Glu	Arg
				125					130					135
Glu	Met	Gly	Leu	Leu	Leu	Ser	Val	Val	Val	Thr	Ser	Arg	Pro	
				140					145					150
Glu	Ala	Phe	Gln	Pro	His	His	Arg	Glu	Leu	Leu	Arg	Leu	Leu	Asn
				155					160					165

Glu Thr Leu Gly	Glu Val Gly Ser Pro	Gly Leu Leu Phe Tyr Ser	170	175	180
Leu Arg Thr Leu	Thr Thr Met Ala Pro	Tyr Leu Ser Thr Glu Asp	185	190	195
Val Pro Leu Ala	Arg Met Leu Val Pro	Lys Leu Ile Met Ala Met	200	205	210
Gln Thr Leu Ile	Pro Ile Asp Glu Ala	Lys Ala Cys Glu Ala Leu	215	220	225
Glu Ala Leu Asp	Glu Leu Leu Glu Ser	Glu Val Pro Val Ile Thr	230	235	240
Pro Tyr Leu Ser	Glu Val Leu Thr Phe	Cys Leu Glu Val Ala Arg	245	250	255
Asn Val Ala Leu	Gly Asn Ala Ile Arg	Ile Arg Ile Leu Cys Cys	260	265	270
Leu Thr Phe Leu	Val Lys Val Lys Ser	Lys Ala Leu Leu Lys Asn	275	280	285
Arg Leu Leu Pro	Pro Leu Leu His Thr	Leu Phe Pro Ile Val Ala	290	295	300
Ala Glu Pro Pro	Pro Gly Gln Leu Asp	Pro Glu Asp Gln Asp Ser	305	310	315
Glu Glu Glu Glu	Leu Glu Ile Glu Leu	Met Gly Glu Thr Pro Lys	320	325	330
His Phe Ala Val	Gln Val Val Asp Met	Leu Ala Leu His Leu Pro	335	340	345
Pro Glu Lys Leu	Cys Pro Gln Leu Met	Pro Met Leu Glu Glu Ala	350	355	360
Leu Arg Ser Glu	Ser Pro Tyr Gln Arg	Lys Ala Gly Leu Leu Val	365	370	375
Leu Ala Val Leu	Ser Asp Gly Ala Gly	Asp His Ile Arg Gln Arg	380	385	390
Leu Leu Pro Pro	Leu Leu Gln Ile Val	Cys Lys Gly Leu Glu Asp	395	400	405
Pro Ser Gln Val	Val Arg Asn Ala Ala	Leu Phe Ala Leu Gly Gln	410	415	420
Phe Ser Glu Asn	Leu Gln Pro His Ile	Ser Ser Tyr Ser Arg Glu	425	430	435
Val Met Pro Leu	Leu Leu Ala Tyr Leu	Lys Ser Val Pro Leu Gly	440	445	450
His Thr His His	Leu Ala Lys Ala Cys	Tyr Ala Leu Glu Asn Phe	455	460	465
Val Glu Asn Leu	Gly Pro Lys Val Gln	Pro Tyr Leu Pro Glu Leu	470	475	480
Met Glu Cys Met	Leu Gln Leu Leu Arg	Asn Pro Ser Ser Pro Arg	485	490	495
Ala Lys Glu Leu	Ala Val Ser Ala Leu	Gly Ala Ile Ala Thr Ala	500	505	510
Ala Gln Ala Ser	Leu Leu Pro Tyr Phe	Pro Ala Ile Met Glu His	515	520	525
Leu Arg Glu Phe	Leu Leu Thr Gly Arg	Glu Asp Leu Gln Pro Val	530	535	540
Gln Ile Gln Ser	Leu Glu Thr Leu Gly	Val Leu Ala Arg Ala Val	545	550	555
Gly Glu Pro Met	Arg Pro Leu Ala Glu	Glu Cys Cys Gln Leu Gly	560	565	570
Leu Gly Leu Cys	Asp Gln Val Asp Asp	Ala Asp Leu Arg Arg Cys	575	580	585
Thr Tyr Ser Leu	Phe Ala Ala Leu Ser	Gly Leu Met Gly Glu Gly	590	595	600
Leu Ala Pro His	Leu Glu Gln Ile Thr	Thr Leu Met Leu Leu Ser	605	610	615
Leu Arg Ser Thr	Glu Gly Ile Val Pro	Gln Tyr Asp Gly Ser Ser	620	625	630
Ser Phe Leu Leu	Phe Asp Asp Glu Ser	Asp Gly Glu Glu Glu Glu			

	635		640		645
Glu Leu Met Asp	Glu Asp Val Glu Glu	Glu Asp Asp Ser Glu	Ile		
	650		655		660
Ser Gly Tyr Ser	Val Glu Asn Ala Phe	Phe Asp Glu Lys Glu	Asp		
	665		670		675
Thr Cys Ala Ala	Val Gly Glu Ile Ser	Val Asn Thr Ser Val	Ala		
	680		685		690
Phe Leu Pro Tyr	Met Glu Ser Val Phe	Glu Glu Val Phe Lys	Leu		
	695		700		705
Leu Glu Cys Pro	His Leu Asn Val Arg	Lys Ala Ala His Glu	Ala		
	710		715		720
Leu Gly Gln Phe	Cys Cys Ala Leu His	Lys Ala Cys Gln Ser	Cys		
	725		730		735
Pro Ser Glu Pro	Asn Thr Ala Ala Leu	Gln Ala Ala Leu Ala	Arg		
	740		745		750
Val Val Pro Ser	Tyr Met Gln Ala Val	Asn Arg Glu Arg Glu	Arg		
	755		760		765
Gln Val Val Met	Ala Val Leu Glu Ala	Leu Thr Gly Val Leu	Arg		
	770		775		780
Ser Cys Gly Thr	Leu Thr Leu Lys Pro	Pro Gly Arg Leu Ala	Glu		
	785		790		795
Leu Cys Gly Val	Leu Lys Ala Val Leu	Gln Arg Lys Thr Ala	Cys		
	800		805		810
Gln Asp Thr Asp	Glu Glu Glu Glu Glu	Glu Asp Asp Asp Gln	Ala		
	815		820		825
Glu Tyr Asp Ala	Met Leu Leu Glu His	Ala Gly Glu Ala Ile	Pro		
	830		835		840
Ala Leu Ala Ala	Ala Ala Gly Gly Asp	Ser Phe Ala Pro Phe	Phe		
	845		850		855
Ala Gly Phe Leu	Pro Leu Leu Val Cys	Lys Thr Lys Gln Gly	Cys		
	860		865		870
Thr Val Ala Glu	Lys Ser Phe Ala Val	Gly Thr Leu Ala Glu	Thr		
	875		880		885
Ile Gln Gly Leu	Gly Ala Ala Ser Ala	Gln Phe Val Ser Arg	Leu		
	890		895		900
Leu Pro Val Leu	Leu Ser Thr Ala Gln	Glu Ala Asp Pro Glu	Val		
	905		910		915
Arg Ser Asn Ala	Ile Phe Gly Met Gly	Val Leu Ala Glu His	Gly		
	920		925		930
Gly His Pro Ala	Gln Glu His Phe Pro	Lys Leu Leu Gly Leu	Leu		
	935		940		945
Phe Pro Leu Leu	Ala Arg Glu Arg His	Asp Arg Val Arg Asp	Asn		
	950		955		960
Ile Cys Gly Ala	Leu Ala Arg Leu Leu	Met Ala Ser Pro Thr	Arg		
	965		970		975
Lys Pro Glu Pro	Gln Val Leu Ala Ala	Leu Leu His Ala Leu	Pro		
	980		985		990
Leu Lys Glu Asp	Leu Glu Glu Trp Val	Thr Ile Gly Arg Leu	Phe		
	995		1000		1005
Ser Phe Leu Tyr	Gln Ser Ser Pro Asp	Gln Val Ile Asp Val	Ala		
	1010		1015		1020
Pro Glu Leu Leu	Arg Ile Cys Ser Leu	Ile Leu Ala Asp Asn	Lys		
	1025		1030		1035
Ile Pro Pro Asp	Thr Lys Ala Ala Leu	Leu Leu Leu Leu Thr	Phe		
	1040		1045		1050
Leu Ala Lys Gln	His Thr Asp Ser Phe	Gln Ala Ala Leu Gly	Ser		
	1055		1060		1065
Leu Pro Val Asp	Lys Ala Gln Glu Leu	Gln Ala Val Leu Gly	Leu		
	1070		1075		1080
Ser					

<210> 3

<211> 1172  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3552033CD1

<400> 3

Met	Tyr	Val	Arg	Val	Ser	Phe	Asp	Thr	Lys	Pro	Asp	Leu	Leu	Leu	1	5	10	15
His	Leu	Met	Thr	Lys	Glu	Trp	Gln	Leu	Glu	Leu	Pro	Lys	Leu	Leu	20	25	30	35
Ile	Ser	Val	His	Gly	Gly	Leu	Gln	Asn	Phe	Glu	Leu	Gln	Pro	Lys	40	45	50	55
Leu	Lys	Gln	Val	Phe	Gly	Lys	Gly	Leu	Ile	Lys	Ala	Ala	Met	Thr	60	65	70	75
Thr	Gly	Ala	Trp	Ile	Phe	Thr	Gly	Gly	Val	Asn	Thr	Gly	Val	Ile	80	85	90	95
Arg	His	Val	Gly	Asp	Ala	Leu	Lys	Asp	His	Ala	Ser	Lys	Ser	Arg	100	105	110	115
Gly	Lys	Ile	Cys	Thr	Ile	Gly	Ile	Ala	Pro	Trp	Gly	Ile	Val	Glu	120	125	130	135
Asn	Gln	Glu	Asp	Leu	Ile	Gly	Arg	Asp	Val	Val	Arg	Pro	Tyr	Gln	140	145	150	155
Thr	Met	Ser	Asn	Pro	Met	Ser	Lys	Leu	Thr	Val	Leu	Asn	Ser	Met	160	165	170	175
His	Ser	His	Phe	Ile	Leu	Ala	Asp	Asn	Gly	Thr	Thr	Gly	Lys	Tyr	180	185	190	195
Gly	Ala	Glu	Val	Lys	Leu	Arg	Arg	Gln	Leu	Glu	Lys	His	Ile	Ser	200	205	210	215
Leu	Gln	Lys	Ile	Asn	Thr	Arg	Ile	Gly	Gln	Gly	Val	Pro	Val	Val	220	225	230	235
Ala	Leu	Ile	Val	Glu	Gly	Gly	Pro	Asn	Val	Ile	Ser	Ile	Val	Leu	240	245	250	255
Glu	Tyr	Leu	Arg	Asp	Thr	Pro	Pro	Val	Pro	Val	Val	Val	Cys	Asp	260	265	270	275
Gly	Ser	Gly	Arg	Ala	Ser	Asp	Ile	Leu	Ala	Phe	Gly	His	Lys	Tyr	280	285	290	295
Ser	Glu	Glu	Gly	Gly	Leu	Ile	Asn	Glu	Ser	Leu	Arg	Asp	Gln	Leu	300	305	310	315
Leu	Val	Thr	Ile	Gln	Lys	Thr	Phe	Thr	Tyr	Thr	Arg	Thr	Gln	Ala	320	325	330	335
Gln	His	Leu	Phe	Ile	Ile	Leu	Met	Glu	Cys	Met	Lys	Lys	Lys	Glu	340	345	350	355
Leu	Ile	Thr	Val	Phe	Arg	Met	Gly	Ser	Glu	Gly	His	Gln	Asp	Ile	360	365	370	375
Asp	Leu	Ala	Ile	Leu	Thr	Ala	Leu	Leu	Lys	Gly	Ala	Asn	Ala	Ser	380	385	390	395
Ala	Pro	Asp	Gln	Leu	Ser	Leu	Ala	Leu	Ala	Trp	Asn	Arg	Val	Asp	400			
Ile	Ala	Arg	Ser	Gln	Ile	Phe	Ile	Tyr	Gly	Gln	Gln	Trp	Pro	Val				
Gly	Ser	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val	Leu	Asp	Arg				
Val	Asp	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Ser	Met	His				
Arg	Phe	Leu	Thr	Ile	Ser	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr	Arg				
His	Gly	Pro	Ser	Asn	Thr	Leu	Tyr	His	Leu	Val	Arg	Asp	Val	Lys				
Lys	Gly	Asn	Leu	Pro	Pro	Asp	Tyr	Arg	Ile	Ser	Leu	Ile	Asp	Ile				

Gly	Leu	Val	Ile	Glu	Tyr	Leu	Met	Gly	Gly	Ala	Tyr	Arg	Cys	Asn	
				410					415					420	
Tyr	Thr	Arg	Lys	Arg	Phe	Arg	Thr	Leu	Tyr	His	Asn	Leu	Phe	Gly	
				425					430					435	
Pro	Lys	Arg	Pro	Lys	Ala	Leu	Lys	Leu	Leu	Gly	Met	Glu	Asp	Asp	
				440					445					450	
Ile	Pro	Leu	Arg	Arg	Gly	Arg	Lys	Thr	Thr	Lys	Lys	Arg	Glu	Glu	
				455					460					465	
Glu	Val	Asp	Ile	Asp	Leu	Asp	Asp	Pro	Glu	Ile	Asn	His	Phe	Pro	
				470					475					480	
Phe	Pro	Phe	His	Glu	Leu	Met	Val	Trp	Ala	Val	Leu	Met	Lys	Arg	
				485					490					495	
Gln	Lys	Met	Ala	Leu	Phe	Phe	Trp	Gln	His	Gly	Glu	Glu	Ala	Met	
				500					505					510	
Ala	Lys	Ala	Leu	Val	Ala	Cys	Lys	Leu	Cys	Lys	Ala	Met	Ala	His	
				515					520					525	
Glu	Ala	Ser	Glu	Asn	Asp	Met	Val	Asp	Asp	Ile	Ser	Gln	Glu	Leu	
				530					535					540	
Asn	His	Asn	Ser	Arg	Asp	Phe	Gly	Gln	Leu	Ala	Val	Glu	Leu	Leu	
				545					550					555	
Asp	Gln	Ser	Tyr	Lys	Gln	Asp	Glu	Gln	Leu	Ala	Met	Lys	Leu	Leu	
				560					565					570	
Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ala	Thr	Cys	Leu	Gln	Leu	
				575					580					585	
Ala	Val	Ala	Ala	Lys	His	Arg	Asp	Phe	Ile	Ala	His	Thr	Cys	Ser	
				590					595					600	
Gln	Met	Leu	Leu	Thr	Asp	Met	Trp	Met	Gly	Arg	Leu	Arg	Met	Arg	
				605					610					615	
Lys	Asn	Ser	Gly	Leu	Lys	Val	Ile	Leu	Gly	Ile	Leu	Leu	Pro	Pro	
				620					625					630	
Ser	Ile	Leu	Ser	Leu	Glu	Phe	Lys	Asn	Lys	Asp	Asp	Met	Pro	Tyr	
				635					640					645	
Met	Ser	Gln	Ala	Gln	Glu	Ile	His	Leu	Gln	Glu	Lys	Glu	Ala	Glu	
				650					655					660	
Glu	Pro	Glu	Lys	Pro	Thr	Lys	Glu	Lys	Glu	Glu	Glu	Asp	Met	Glu	
				665					670					675	
Leu	Ile	Ala	Met	Leu	Gly	Arg	Asn	Asn	Gly	Glu	Ser	Ser	Arg	Lys	
				680					685					690	
Lys	Asp	Glu	Glu	Glu	Val	Gln	Ser	Glu	His	Arg	Leu	Ile	Pro	Leu	
				695					700					705	
Gly	Arg	Lys	Ile	Tyr	Glu	Phe	Tyr	Asn	Ala	Pro	Ile	Val	Lys	Phe	
				710					715					720	
Trp	Phe	Tyr	Thr	Leu	Ala	Tyr	Ile	Gly	Tyr	Leu	Met	Leu	Phe	Asn	
				725					730					735	
Tyr	Ile	Val	Leu	Val	Lys	Met	Glu	Arg	Trp	Pro	Pro	Thr	Gln	Glu	
				740					745					750	
Trp	Ile	Val	Ile	Ser	Tyr	Ile	Phe	Thr	Leu	Gly	Ile	Glu	Lys	Met	
				755					760					765	
Arg	Glu	Ile	Leu	Met	Ser	Glu	Pro	Gly	Lys	Leu	Leu	Gln	Lys	Val	
				770					775					780	
Lys	Val	Trp	Leu	Gln	Glu	His	Trp	Asn	Val	Thr	Asp	Leu	Ile	Ala	
				785					790					795	
Ile	Leu	Leu	Phe	Ser	Val	Gly	Met	Ile	Leu	Arg	Leu	Gln	Asp	Gln	
				800					805					810	
Pro	Phe	Arg	Ser	Asp	Gly	Arg	Val	Ile	Tyr	Cys	Val	Asn	Ile	Ile	
				815					820					825	
Tyr	Trp	Tyr	Ile	Arg	Leu	Leu	Asp	Ile	Phe	Gly	Val	Asn	Lys	Tyr	
				830					835					840	
Leu	Gly	Pro	Tyr	Val	Met	Met	Ile	Gly	Lys	Met	Met	Ile	Asp	Met	
				845					850					855	
Met	Tyr	Phe	Val	Ile	Ile	Met	Leu	Val	Val	Leu	Met	Ser	Phe	Gly	
				860					865					870	
Val	Ala	Arg	Gln	Ala	Ile	Leu	Phe	Pro	Asn	Glu	Glu	Pro	Ser	Trp	

	875		880		885
Lys Leu Ala Lys	Asn Ile Phe Tyr Met	Pro Tyr Trp Met Ile Tyr			
	890		895		900
Gly Glu Val Phe	Ala Asp Gln Ile Asp	Pro Pro Cys Gly Gln Asn			
	905		910		915
Glu Thr Arg Glu	Asp Gly Lys Ile Ile	Gln Leu Pro Pro Cys Lys			
	920		925		930
Thr Gly Ala Trp	Ile Val Pro Ala Ile	Met Ala Cys Tyr Leu Leu			
	935		940		945
Val Ala Asn Ile	Leu Leu Val Asn Leu	Leu Ile Ala Val Phe Asn			
	950		955		960
Asn Thr Phe Phe	Glu Val Lys Ser Ile	Ser Asn Gln Val Trp Lys			
	965		970		975
Phe Gln Arg Tyr	Gln Leu Ile Met Thr	Phe His Glu Arg Pro Val			
	980		985		990
Leu Pro Pro Pro	Leu Ile Ile Phe Ser	His Met Thr Met Ile Phe			
	995		1000		1005
Gln His Leu Cys	Cys Arg Trp Arg Lys	His Glu Ser Asp Pro Asp			
	1010		1015		1020
Glu Arg Asp Tyr	Gly Leu Lys Leu Phe	Ile Thr Asp Asp Glu Leu			
	1025		1030		1035
Lys Lys Val His	Asp Phe Glu Glu Gln	Cys Ile Glu Glu Tyr Phe			
	1040		1045		1050
Arg Glu Lys Asp	Asp Arg Phe Asn Ser	Ser Asn Asp Glu Arg Ile			
	1055		1060		1065
Arg Val Thr Ser	Glu Arg Val Glu Asn	Met Ser Met Arg Leu Glu			
	1070		1075		1080
Glu Val Asn Glu	Arg Glu His Ser Met	Lys Ala Ser Leu Gln Thr			
	1085		1090		1095
Val Asp Ile Arg	Leu Ala Gln Leu Glu	Asp Leu Ile Gly Arg Met			
	1100		1105		1110
Ala Thr Ala Leu	Glu Arg Leu Thr Gly	Leu Glu Arg Ala Glu Ser			
	1115		1120		1125
Asn Lys Ile Arg	Ser Arg Thr Ser Ser	Asp Cys Thr Asp Ala Ala			
	1130		1135		1140
Tyr Ile Val Arg	Gln Ser Ser Phe Asn	Ser Gln Glu Gly Asn Thr			
	1145		1150		1155
Phe Lys Leu Gln	Glu Ser Ile Asp Pro	Ala Glu His Pro Leu Tyr			
	1160		1165		1170
Ser Val					

<210> 4

<211> 742

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4778139CD1

<400> 4

Met Glu Ile Ala Gly	Glu Lys Lys Tyr Gln	Tyr Tyr Ser Met Pro
1	5	10
Ser Ser Leu Leu Arg	Ser Leu Ala Gly	Gly Arg His Pro Pro Arg
	20	25
Val Gln Arg Gly Ala	Leu Leu Glu Ile Thr	Asn Ser Lys Arg Glu
	35	40
Ala Thr Asn Val Arg	Asn Asp Gln Glu Arg	Gln Glu Thr Gln Ser
	50	55
Ser Ile Val Val Ser	Gly Val Ser Pro Asn	Arg Gln Ala His Ser
	65	70
Lys Tyr Gly Gln Phe	Leu Leu Val Pro Ser	Asn Leu Lys Arg Val

Pro Phe Ser Ala	80	85	90
Glu Thr Val Leu Pro	95	100	105
Pro Asp Val Leu Ala Thr Glu Gln Asp	110	115	120
Val His Leu Thr Gly Gln Thr Pro Val	125	130	135
Val Asp Glu Pro Lys Glu Lys His Glu	140	145	150
Asp Gly Gln Asp Lys Ala Glu Gln Val	155	160	165
Thr Pro Glu Pro Met Thr Met Ser Ser	170	175	180
Ile Met Gln Thr Gly Gly Tyr Val	185	190	195
Gly Asp Glu Thr Gly Val Phe Ser Ile	200	205	210
Cys Leu Pro Ser Thr Cys Asp Ser Asp	215	220	225
His Gln Lys Gln Ala Gln Asp Ser Ser	230	235	240
Arg Ser Ala Gln Ser Ser Glu Cys Ser	245	250	255
Gln Pro Asn Thr Ser Phe Trp Ile Asn	260	265	270
Pro Phe Ala Arg Ser His Ser Phe Arg	275	280	285
Leu Met Lys Ile Cys Lys Ile Lys Asn	290	295	300
Ile Gly Gln Gly Ala Trp Val Lys Ala	305	310	315
Arg Arg Leu Ser Lys Lys Lys Lys Asn	320	325	330
Pro Ile Ile Thr Val Asn Ala Cys Ser	335	340	345
Pro Glu Pro Gly Glu Asn Ser Ile Ser	350	355	360
Asn Trp Phe Thr Val Ser Lys Phe Ser	365	370	375
Tyr Ile His Gln Lys Met Lys Thr Lys	380	385	390
Ile Gln Ile Ser Asp Tyr Leu Lys Gln	395	400	405
Asp Leu Ser Lys Asn Ser Leu Trp Asn	410	415	420
Asn Arg Asn Ser Leu Ser Ser Leu Ile	425	430	435
Leu Lys Ser Pro Gln Glu Pro His His	440	445	450
Leu Phe Ala Ala Gly Glu Glu Ile Thr	455	460	465
Ser Ser Pro Leu Asn Leu Asp Lys Ser	470	475	480
Arg Gly Arg Ala Ala Met Ile Gln Val	485	490	495
Asp Gly Gly Leu Arg Lys Ala Met Arg	500	505	510
Glu Asp Asp Ile Leu Lys Pro Gly Gln	515	520	525
Phe Leu Pro Glu Val Val Arg Thr Trp	530	535	540
Ser Thr Val Leu His Leu Cys Leu Arg	545	550	555

Ala	Ala	Gln	Lys	Leu	Ile	Tyr	Thr	Phe	Asn	Gln	Val	Lys	Pro	Gln	
				560					565					570	
Thr	Ile	Pro	Tyr	Thr	Pro	Arg	Phe	Leu	Glu	Val	Phe	Leu	Ile	Tyr	
				575					580					585	
Cys	His	Ser	Ala	Asn	Gln	Trp	Leu	Thr	Ile	Glu	Lys	Tyr	Met	Thr	
				590					595					600	
Gly	Glu	Phe	Arg	Lys	Tyr	Asn	Asn	Asn	Asn	Gly	Asp	Glu	Ile	Thr	
				605					610					615	
Pro	Thr	Asn	Thr	Leu	Glu	Glu	Leu	Met	Leu	Ala	Phe	Ser	His	Trp	
				620					625					630	
Thr	Tyr	Glu	Tyr	Thr	Arg	Gly	Glu	Leu	Leu	Val	Leu	Asp	Leu	Gln	
				635					640					645	
Gly	Val	Gly	Glu	Asn	Leu	Thr	Asp	Pro	Ser	Val	Ile	Lys	Pro	Glu	
				650					655					660	
Val	Lys	Gln	Ser	Arg	Gly	Met	Val	Phe	Gly	Pro	Ala	Asn	Leu	Gly	
				665					670					675	
Glu	Asp	Ala	Ile	Arg	Asn	Phe	Ile	Ala	Lys	His	His	Trp	Asn	Ser	
				680					685					690	
Cys	Cys	Arg	Lys	Leu	Lys	Leu	Pro	Asp	Leu	Lys	Arg	Asn	Asp	Tyr	
				695					700					705	
Ser	Pro	Glu	Arg	Ile	Asn	Ser	Thr	Phe	Gly	Leu	Glu	Ile	Lys	Ile	
				710					715					720	
Glu	Ser	Ala	Glu	Glu	Pro	Pro	Ala	Arg	Glu	Thr	Gly	Arg	Asn	Ser	
				725					730					735	
Pro	Glu	Asp	Asp	Met	Gln	Leu									
				740											

<210> 5

<211> 577

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4787433CD1

<400> 5

Met	Gly	Ser	Arg	His	Phe	Glu	Gly	Ile	Tyr	Asp	His	Val	Gly	His	
1				5					10					15	
Phe	Gly	Arg	Phe	Gln	Arg	Val	Leu	Tyr	Phe	Ile	Cys	Ala	Phe	Gln	
				20					25					30	
Asn	Ile	Ser	Cys	Gly	Ile	His	Tyr	Leu	Ala	Ser	Val	Phe	Met	Gly	
				35					40					45	
Val	Thr	Pro	His	His	Val	Cys	Arg	Pro	Pro	Gly	Asn	Val	Ser	Gln	
				50					55					60	
Val	Val	Phe	His	Asn	His	Ser	Asn	Trp	Ser	Leu	Glu	Asp	Thr	Gly	
				65					70					75	
Ala	Leu	Leu	Ser	Ser	Gly	Gln	Lys	Asp	Tyr	Val	Thr	Val	Gln	Leu	
				80					85					90	
Gln	Asn	Gly	Glu	Ile	Trp	Glu	Leu	Ser	Arg	Cys	Ser	Arg	Asn	Lys	
				95					100					105	
Arg	Glu	Asn	Thr	Ser	Ser	Leu	Gly	Tyr	Glu	Tyr	Thr	Gly	Ser	Lys	
				110					115					120	
Lys	Glu	Phe	Pro	Cys	Val	Asp	Gly	Tyr	Ile	Tyr	Asp	Gln	Asn	Thr	
				125					130					135	
Trp	Lys	Ser	Thr	Ala	Val	Thr	Gln	Trp	Asn	Leu	Val	Cys	Asp	Arg	
				140					145					150	
Lys	Trp	Leu	Ala	Met	Leu	Ile	Gln	Pro	Leu	Phe	Met	Phe	Gly	Val	
				155					160					165	
Leu	Leu	Gly	Ser	Val	Thr	Phe	Gly	Tyr	Phe	Ser	Asp	Arg	Leu	Gly	
				170					175					180	
Arg	Arg	Val	Val	Leu	Trp	Ala	Thr	Ser	Ser	Ser	Met	Phe	Leu	Phe	
				185					190					195	



Gly	Ile	Ala	Ala	Ala	Phe	Ala	Val	Asp	Tyr	Tyr	Thr	Phe	Met	Ala	
				200					205					210	
Ala	Arg	Phe	Phe	Leu	Ala	Met	Val	Ala	Ser	Gly	Tyr	Leu	Val	Val	
				215					220					225	
Gly	Phe	Val	Tyr	Val	Met	Glu	Phe	Ile	Gly	Met	Lys	Ser	Arg	Thr	
				230					235					240	
Trp	Ala	Ser	Val	His	Leu	His	Ser	Phe	Phe	Ala	Val	Gly	Thr	Leu	
				245					250					255	
Leu	Val	Ala	Leu	Thr	Gly	Tyr	Leu	Val	Arg	Thr	Trp	Trp	Leu	Tyr	
				260					265					270	
Gln	Met	Ile	Leu	Ser	Thr	Val	Thr	Val	Pro	Phe	Ile	Leu	Cys	Cys	
				275					280					285	
Trp	Val	Leu	Pro	Glu	Thr	Pro	Phe	Trp	Leu	Leu	Ser	Glu	Gly	Arg	
				290					295					300	
Tyr	Glu	Glu	Ala	Gln	Lys	Ile	Val	Asp	Ile	Met	Ala	Lys	Trp	Asn	
				305					310					315	
Arg	Ala	Ser	Ser	Cys	Lys	Leu	Ser	Glu	Leu	Leu	Ser	Leu	Asp	Leu	
				320					325					330	
Gln	Gly	Pro	Val	Ser	Asn	Ser	Pro	Thr	Glu	Val	Gln	Lys	His	Asn	
				335					340					345	
Leu	Ser	Tyr	Leu	Phe	Tyr	Asn	Trp	Ser	Ile	Thr	Lys	Arg	Thr	Leu	
				350					355					360	
Thr	Val	Trp	Leu	Ile	Trp	Phe	Thr	Gly	Ser	Leu	Gly	Phe	Tyr	Ser	
				365					370					375	
Phe	Ser	Leu	Asn	Ser	Val	Asn	Leu	Gly	Gly	Asn	Glu	Tyr	Leu	Asn	
				380					385					390	
Leu	Phe	Leu	Leu	Gly	Val	Val	Glu	Ile	Pro	Ala	Tyr	Thr	Phe	Val	
				395					400					405	
Cys	Ile	Ala	Met	Asp	Lys	Val	Gly	Arg	Arg	Thr	Val	Leu	Ala	Tyr	
				410					415					420	
Ser	Leu	Phe	Cys	Ser	Ala	Leu	Ala	Cys	Gly	Val	Val	Met	Val	Ile	
				425					430					435	
Pro	Gln	Lys	His	Tyr	Ile	Leu	Gly	Val	Val	Thr	Ala	Met	Val	Gly	
				440					445					450	
Lys	Phe	Ala	Ile	Gly	Ala	Ala	Phe	Gly	Leu	Ile	Tyr	Leu	Tyr	Thr	
				455					460					465	
Ala	Glu	Leu	Tyr	Pro	Thr	Ile	Val	Arg	Ser	Leu	Ala	Val	Gly	Ser	
				470					475					480	
Gly	Ser	Met	Val	Cys	Arg	Leu	Ala	Ser	Ile	Leu	Ala	Pro	Phe	Ser	
				485					490					495	
Val	Asp	Leu	Ser	Ser	Ile	Trp	Ile	Phe	Ile	Pro	Gln	Leu	Phe	Val	
				500					505					510	
Gly	Thr	Met	Ala	Leu	Leu	Ser	Gly	Val	Leu	Thr	Leu	Lys	Leu	Pro	
				515					520					525	
Glu	Thr	Leu	Gly	Lys	Arg	Leu	Ala	Thr	Thr	Trp	Glu	Glu	Ala	Ala	
				530					535					540	
Lys	Leu	Glu	Ser	Glu	Asn	Glu	Ser	Lys	Ser	Ser	Lys	Leu	Leu	Leu	
				545					550					555	
Thr	Thr	Asn	Asn	Ser	Gly	Leu	Glu	Lys	Thr	Glu	Ala	Ile	Thr	Pro	
				560					565					570	
Arg	Asp	Ser	Gly	Leu	Gly	Glu									
				575											

<210> 6

<211> 462

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7483598CD1

<400> 6

Met	Gly	Tyr	Gln	Arg	Gln	Glu	Pro	Val	Ile	Pro	Pro	Gln	Arg	Asp
1				5					10					15
Leu	Asp	Asp	Arg	Glu	Thr	Leu	Val	Ser	Glu	His	Glu	Tyr	Lys	Glu
				20					25					30
Lys	Thr	Cys	Gln	Ser	Ala	Ala	Leu	Phe	Asn	Val	Val	Asn	Ser	Ile
				35					40					45
Ile	Gly	Ser	Gly	Ile	Ile	Gly	Leu	Pro	Tyr	Ser	Met	Lys	Gln	Ala
				50					55					60
Gly	Phe	Pro	Leu	Gly	Ile	Leu	Leu	Leu	Phe	Trp	Val	Ser	Tyr	Val
				65					70					75
Thr	Asp	Phe	Ser	Leu	Val	Leu	Leu	Ile	Lys	Gly	Gly	Ala	Leu	Ser
				80					85					90
Gly	Thr	Asp	Thr	Tyr	Gln	Ser	Leu	Val	Asn	Lys	Thr	Phe	Gly	Phe
				95					100					105
Pro	Gly	Tyr	Leu	Leu	Leu	Ser	Val	Leu	Gln	Phe	Leu	Tyr	Pro	Phe
				110					115					120
Ile	Ala	Met	Ile	Ser	Tyr	Asn	Ile	Ile	Ala	Gly	Asp	Thr	Leu	Ser
				125					130					135
Lys	Val	Phe	Gln	Arg	Ile	Pro	Gly	Val	Asp	Pro	Glu	Asn	Val	Phe
				140					145					150
Ile	Gly	Arg	His	Phe	Ile	Ile	Gly	Leu	Ser	Thr	Val	Thr	Phe	Thr
				155					160					165
Leu	Pro	Leu	Ser	Leu	Tyr	Arg	Asn	Ile	Ala	Lys	Leu	Gly	Lys	Val
				170					175					180
Ser	Leu	Ile	Ser	Thr	Gly	Leu	Thr	Thr	Leu	Ile	Leu	Gly	Ile	Val
				185					190					195
Met	Ala	Arg	Ala	Ile	Ser	Leu	Gly	Pro	His	Ile	Pro	Lys	Thr	Glu
				200					205					210
Asp	Ala	Trp	Val	Phe	Ala	Lys	Pro	Asn	Ala	Ile	Gln	Ala	Val	Gly
				215					220					225
Val	Met	Ser	Phe	Ala	Phe	Ile	Cys	His	His	Asn	Ser	Phe	Leu	Val
				230					235					240
Tyr	Ser	Ser	Leu	Glu	Glu	Pro	Thr	Val	Ala	Lys	Trp	Ser	Arg	Leu
				245					250					255
Ile	His	Met	Ser	Ile	Val	Ile	Ser	Val	Phe	Ile	Cys	Ile	Phe	Phe
				260					265					270
Ala	Thr	Cys	Gly	Tyr	Leu	Thr	Phe	Thr	Gly	Phe	Thr	Gln	Gly	Asp
				275					280					285
Leu	Phe	Glu	Asn	Tyr	Cys	Arg	Asn	Asp	Asp	Leu	Val	Thr	Phe	Gly
				290					295					300
Arg	Phe	Cys	Tyr	Gly	Val	Thr	Val	Ile	Leu	Thr	Tyr	Pro	Met	Glu
				305					310					315
Cys	Phe	Val	Thr	Arg	Glu	Val	Ile	Ala	Asn	Val	Phe	Phe	Gly	Gly
				320					325					330
Asn	Leu	Ser	Ser	Val	Phe	His	Ile	Val	Val	Thr	Val	Met	Val	Ile
				335					340					345
Thr	Val	Ala	Thr	Leu	Val	Ser	Leu	Leu	Ile	Asp	Cys	Leu	Gly	Ile
				350					355					360
Val	Leu	Glu	Leu	Asn	Gly	Val	Leu	Cys	Ala	Thr	Pro	Leu	Ile	Phe
				365					370					375
Ile	Ile	Pro	Ser	Ala	Cys	Tyr	Leu	Lys	Leu	Ser	Glu	Glu	Pro	Arg
				380					385					390
Thr	His	Ser	Asp	Lys	Ile	Met	Ser	Cys	Val	Met	Leu	Pro	Ile	Gly
				395					400					405
Ala	Val	Val	Met	Val	Phe	Gly	Phe	Val	Met	Ala	Ile	Thr	Asn	Thr
				410					415					420
Gln	Asp	Cys	Thr	His	Gly	Gln	Glu	Met	Phe	Tyr	Cys	Phe	Pro	Asp
				425					430					435
Asn	Phe	Ser	Leu	Thr	Asn	Thr	Ser	Glu	Ser	His	Val	Gln	Gln	Thr
				440					445					450
Thr	Gln	Leu	Ser	Thr	Leu	Asn	Ile	Ser	Ile	Phe	Gln			
				455					460					

<210> 7  
 <211> 1018  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7484823CD1

<400> 7  
 Met Gly Glu Gly Asp Ser Gly Leu Arg Ser Pro Arg Gly Pro Gln  
 1 5 10 15  
 Glu Gly Gly Lys Cys Ser Arg Glu Lys Gln Lys Arg Asn Met Glu  
 20 25 30  
 Glu Leu Lys Lys Glu Val Val Met Asp Asp His Lys Leu Thr Leu  
 35 40 45  
 Glu Glu Leu Ser Thr Lys Tyr Ser Val Asp Leu Thr Lys Gly His  
 50 55 60  
 Ser His Gln Arg Ala Lys Glu Ile Leu Thr Arg Gly Gly Pro Asn  
 65 70 75  
 Thr Val Thr Pro Pro Pro Thr Thr Pro Glu Trp Val Lys Phe Cys  
 80 85 90  
 Lys Gln Leu Phe Gly Gly Phe Ser Leu Leu Leu Trp Thr Gly Ala  
 95 100 105  
 Ile Leu Cys Phe Val Ala Tyr Ser Ile Gln Ile Tyr Phe Asn Glu  
 110 115 120  
 Glu Pro Thr Lys Asp Asn Leu Tyr Leu Ser Ile Val Leu Ser Val  
 125 130 135  
 Val Val Ile Val Thr Gly Cys Phe Ser Tyr Tyr Gln Glu Ala Lys  
 140 145 150  
 Ser Ser Lys Ile Met Glu Ser Phe Lys Asn Met Val Pro Gln Gln  
 155 160 165  
 Ala Leu Val Ile Arg Gly Gly Glu Lys Met Gln Ile Asn Val Gln  
 170 175 180  
 Glu Val Val Leu Gly Asp Leu Val Glu Ile Lys Gly Gly Asp Arg  
 185 190 195  
 Val Pro Ala Asp Leu Arg Leu Ile Ser Ala Gln Gly Cys Lys Val  
 200 205 210  
 Asp Asn Ser Ser Leu Thr Gly Glu Ser Glu Pro Gln Ser Arg Ser  
 215 220 225  
 Pro Asp Phe Thr His Glu Asn Pro Leu Glu Thr Arg Asn Ile Cys  
 230 235 240  
 Phe Phe Ser Thr Asn Cys Val Glu Gly Thr Ala Arg Gly Ile Val  
 245 250 255  
 Ile Ala Thr Gly Asp Ser Thr Val Met Gly Arg Ile Ala Ser Leu  
 260 265 270  
 Thr Ser Gly Leu Ala Val Gly Gln Thr Pro Ile Ala Ala Glu Ile  
 275 280 285  
 Glu His Phe Ile His Leu Ile Thr Val Val Ala Val Phe Leu Gly  
 290 295 300  
 Val Thr Phe Phe Ala Leu Ser Leu Leu Leu Gly Tyr Gly Trp Leu  
 305 310 315  
 Glu Ala Ile Ile Phe Leu Ile Gly Ile Ile Val Ala Asn Val Pro  
 320 325 330  
 Glu Gly Leu Leu Ala Thr Val Thr Val Cys Leu Thr Leu Thr Ala  
 335 340 345  
 Lys Arg Met Ala Arg Lys Asn Cys Leu Val Lys Asn Leu Glu Ala  
 350 355 360  
 Val Glu Thr Leu Gly Ser Thr Ser Thr Ile Cys Ser Asp Lys Thr  
 365 370 375  
 Gly Thr Leu Thr Gln Asn Arg Met Thr Val Ala His Met Trp Phe  
 380 385 390  
 Asp Met Thr Val Tyr Glu Ala Asp Thr Thr Glu Glu Gln Thr Gly

				395					400					405
Lys	Thr	Phe	Thr	Lys	Ser	Ser	Asp	Thr	Trp	Phe	Met	Leu	Ala	Arg
				410					415					420
Ile	Ala	Gly	Leu	Cys	Asn	Arg	Ala	Asp	Phe	Lys	Ala	Asn	Gln	Glu
				425					430					435
Ile	Leu	Pro	Ile	Ala	Lys	Arg	Ala	Thr	Thr	Gly	Asp	Ala	Ser	Glu
				440					445					450
Ser	Ala	Leu	Leu	Lys	Phe	Ile	Glu	Gln	Ser	Tyr	Ser	Ser	Val	Ala
				455					460					465
Glu	Met	Arg	Glu	Lys	Asn	Pro	Lys	Val	Ala	Glu	Val	Pro	Phe	Asn
				470					475					480
Ser	Thr	Asn	Lys	Tyr	Gln	Met	Ser	Ile	His	Leu	Arg	Glu	Asp	Ser
				485					490					495
Ser	Gln	Thr	His	Val	Leu	Met	Met	Lys	Gly	Ala	Pro	Glu	Arg	Ile
				500					505					510
Leu	Glu	Phe	Cys	Ser	Thr	Phe	Leu	Leu	Asn	Gly	Gln	Glu	Tyr	Ser
				515					520					525
Met	Asn	Asp	Glu	Met	Lys	Glu	Ala	Phe	Gln	Asn	Ala	Tyr	Leu	Glu
				530					535					540
Leu	Gly	Gly	Leu	Gly	Glu	Arg	Val	Leu	Gly	Phe	Cys	Phe	Leu	Asn
				545					550					555
Leu	Pro	Ser	Ser	Phe	Ser	Lys	Gly	Phe	Pro	Phe	Asn	Thr	Asp	Glu
				560					565					570
Ile	Asn	Phe	Pro	Met	Asp	Asn	Leu	Cys	Phe	Val	Gly	Leu	Ile	Ser
				575					580					585
Met	Ile	Asp	Pro	Pro	Arg	Ala	Ala	Val	Pro	Asp	Ala	Val	Ser	Lys
				590					595					600
Cys	Arg	Ser	Ala	Gly	Ile	Lys	Val	Ile	Met	Val	Thr	Gly	Asp	His
				605					610					615
Pro	Ile	Thr	Ala	Lys	Ala	Ile	Ala	Lys	Gly	Val	Gly	Ile	Ile	Ser
				620					625					630
Glu	Gly	Thr	Glu	Thr	Ala	Glu	Glu	Val	Ala	Ala	Arg	Leu	Lys	Ile
				635					640					645
Pro	Ile	Ser	Lys	Val	Asp	Ala	Ser	Ala	Ala	Lys	Ala	Ile	Val	Val
				650					655					660
His	Gly	Ala	Glu	Leu	Lys	Asp	Ile	Gln	Ser	Lys	Gln	Leu	Asp	Gln
				665					670					675
Ile	Leu	Gln	Asn	His	Pro	Glu	Ile	Val	Phe	Ala	Arg	Thr	Ser	Pro
				680					685					690
Gln	Gln	Lys	Leu	Ile	Ile	Val	Glu	Gly	Cys	Gln	Arg	Leu	Gly	Ala
				695					700					705
Val	Val	Ala	Val	Thr	Gly	Asp	Gly	Val	Asn	Asp	Ser	Pro	Ala	Leu
				710					715					720
Lys	Lys	Ala	Asp	Ile	Gly	Ile	Ala	Met	Gly	Ile	Ser	Gly	Ser	Asp
				725					730					735
Val	Ser	Lys	Gln	Ala	Ala	Asp	Met	Ile	Leu	Leu	Asp	Asp	Asn	Phe
				740					745					750
Ala	Ser	Ile	Val	Thr	Gly	Val	Glu	Glu	Gly	Arg	Leu	Ile	Phe	Asp
				755					760					765
Asn	Leu	Lys	Lys	Ser	Ile	Met	Tyr	Thr	Leu	Thr	Ser	Asn	Ile	Pro
				770					775					780
Glu	Ile	Thr	Pro	Phe	Leu	Met	Phe	Ile	Ile	Leu	Gly	Ile	Pro	Leu
				785					790					795
Pro	Leu	Gly	Thr	Ile	Thr	Ile	Leu	Cys	Ile	Asp	Leu	Gly	Thr	Asp
				800					805					810
Met	Val	Pro	Ala	Ile	Ser	Leu	Ala	Tyr	Glu	Ser	Ala	Glu	Ser	Asp
				815					820					825
Ile	Met	Lys	Arg	Leu	Pro	Arg	Asn	Pro	Lys	Thr	Asp	Asn	Leu	Val
				830					835					840
Asn	His	Arg	Leu	Ile	Gly	Met	Ala	Tyr	Gly	Gln	Ile	Gly	Met	Ile
				845					850					855
Gln	Ala	Leu	Ala	Gly	Phe	Phe	Thr	Tyr	Phe	Val	Ile	Leu	Ala	Glu
				860					865					870

Asn	Gly	Phe	Arg	Pro	Val	Asp	Leu	Leu	Gly	Ile	Arg	Leu	His	Trp	
				875					880					885	
Glu	Asp	Lys	Tyr	Leu	Asn	Asp	Leu	Glu	Asp	Ser	Tyr	Gly	Gln	Gln	
				890					895					900	
Trp	Thr	Tyr	Glu	Gln	Arg	Lys	Val	Val	Glu	Phe	Thr	Cys	Gln	Thr	
				905					910					915	
Ala	Phe	Phe	Val	Thr	Ile	Val	Val	Val	Gln	Trp	Ala	Asp	Leu	Ile	
				920					925					930	
Ile	Ser	Lys	Thr	Arg	Arg	Asn	Ser	Leu	Phe	Gln	Gln	Gly	Met	Arg	
				935					940					945	
Asn	Lys	Val	Leu	Ile	Phe	Gly	Ile	Leu	Glu	Glu	Thr	Leu	Leu	Ala	
				950					955					960	
Ala	Phe	Leu	Ser	Tyr	Thr	Pro	Gly	Met	Asp	Val	Ala	Leu	Arg	Met	
				965					970					975	
Tyr	Pro	Leu	Lys	Ile	Thr	Trp	Trp	Leu	Cys	Ala	Ile	Pro	Tyr	Ser	
				980					985					990	
Ile	Leu	Ile	Phe	Val	Tyr	Asp	Glu	Ile	Arg	Lys	Leu	Leu	Ile	Arg	
				995					1000					1005	
Gln	His	Pro	Asp	Gly	Trp	Val	Glu	Arg	Glu	Thr	Tyr	Tyr			
				1010					1015						

<210> 8

<211> 313

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 143935CD1

<400> 8

Met	Glu	Glu	Met	Ser	Gly	Glu	Ser	Val	Val	Ser	Ser	Ala	Val	Pro	
1				5					10					15	
Ala	Ala	Ala	Thr	Arg	Thr	Thr	Ser	Phe	Lys	Gly	Thr	Ser	Pro	Ser	
				20					25					30	
Ser	Lys	Tyr	Val	Lys	Leu	Asn	Val	Gly	Gly	Ala	Leu	Tyr	Tyr	Thr	
				35					40					45	
Thr	Met	Gln	Thr	Leu	Thr	Lys	Gln	Asp	Thr	Met	Leu	Lys	Ala	Met	
				50					55					60	
Phe	Ser	Gly	Arg	Met	Glu	Val	Leu	Thr	Asp	Ser	Glu	Gly	Trp	Ile	
				65					70					75	
Leu	Ile	Asp	Arg	Cys	Gly	Lys	His	Phe	Gly	Thr	Ile	Leu	Asn	Tyr	
				80					85					90	
Leu	Arg	Asp	Gly	Ala	Val	Pro	Leu	Pro	Glu	Ser	Arg	Arg	Glu	Ile	
				95					100					105	
Glu	Glu	Leu	Leu	Ala	Glu	Ala	Lys	Tyr	Tyr	Leu	Val	Gln	Gly	Leu	
				110					115					120	
Val	Glu	Glu	Cys	Gln	Ala	Ala	Leu	Gln	Asn	Lys	Asp	Thr	Tyr	Glu	
				125					130					135	
Pro	Phe	Cys	Lys	Val	Pro	Val	Ile	Thr	Ser	Ser	Lys	Glu	Glu	Gln	
				140					145					150	
Lys	Leu	Ile	Ala	Thr	Ser	Asn	Lys	Pro	Ala	Val	Lys	Leu	Leu	Tyr	
				155					160					165	
Asn	Arg	Ser	Asn	Asn	Lys	Tyr	Ser	Tyr	Thr	Ser	Asn	Ser	Asp	Asp	
				170					175					180	
Asn	Met	Leu	Lys	Asn	Ile	Glu	Leu	Phe	Asp	Lys	Leu	Ser	Leu	Arg	
				185					190					195	
Phe	Asn	Gly	Arg	Val	Leu	Phe	Ile	Lys	Asp	Val	Ile	Gly	Asp	Glu	
				200					205					210	
Ile	Cys	Cys	Trp	Ser	Phe	Tyr	Gly	Gln	Gly	Arg	Lys	Ile	Ala	Glu	
				215					220					225	
Val	Cys	Cys	Thr	Ser	Ile	Val	Tyr	Ala	Thr	Glu	Lys	Lys	Gln	Thr	
				230					235					240	

Lys	Val	Glu	Phe	Pro	Glu	Ala	Arg	Ile	Tyr	Glu	Glu	Thr	Leu	Asn
				245					250					255
Ile	Leu	Leu	Tyr	Glu	Ala	Gln	Asp	Gly	Arg	Gly	Pro	Asp	Asn	Ala
				260					265					270
Leu	Leu	Glu	Ala	Thr	Gly	Gly	Ala	Ala	Gly	Arg	Ser	His	His	Leu
				275					280					285
Asp	Glu	Asp	Glu	Glu	Arg	Glu	Arg	Ile	Glu	Arg	Val	Arg	Arg	Ile
				290					295					300
His	Ile	Lys	Arg	Pro	Asp	Asp	Arg	Ala	His	Leu	His	Gln		
				305					310					

<210> 9

<211> 921

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5923789CD1

<400> 9

Met	Ala	Trp	Leu	Arg	Leu	Gln	Pro	Leu	Thr	Ser	Ala	Phe	Leu	His
1				5					10					15
Phe	Gly	Leu	Val	Thr	Phe	Val	Leu	Phe	Leu	Asn	Gly	Leu	Arg	Ala
				20					25					30
Glu	Ala	Gly	Gly	Ser	Gly	Asp	Val	Pro	Ser	Thr	Gly	Gln	Asn	Asn
				35					40					45
Glu	Ser	Cys	Ser	Gly	Ser	Ser	Asp	Cys	Lys	Glu	Gly	Val	Ile	Leu
				50					55					60
Pro	Ile	Trp	Tyr	Pro	Glu	Asn	Pro	Ser	Leu	Gly	Asp	Lys	Ile	Ala
				65					70					75
Arg	Val	Ile	Val	Tyr	Phe	Val	Ala	Leu	Ile	Tyr	Met	Phe	Leu	Gly
				80					85					90
Val	Ser	Ile	Ile	Ala	Asp	Arg	Phe	Met	Ala	Ser	Ile	Glu	Val	Ile
				95					100					105
Thr	Ser	Gln	Glu	Arg	Glu	Val	Thr	Ile	Lys	Lys	Pro	Asn	Gly	Glu
				110					115					120
Thr	Ser	Thr	Thr	Thr	Ile	Arg	Val	Trp	Asn	Glu	Thr	Val	Ser	Asn
				125					130					135
Leu	Thr	Leu	Met	Ala	Leu	Gly	Ser	Ser	Ala	Pro	Glu	Ile	Leu	Leu
				140					145					150
Ser	Leu	Ile	Glu	Val	Cys	Gly	His	Gly	Phe	Ile	Ala	Gly	Asp	Leu
				155					160					165
Gly	Pro	Ser	Thr	Ile	Val	Gly	Ser	Ala	Ala	Phe	Asn	Met	Phe	Ile
				170					175					180
Ile	Ile	Gly	Ile	Cys	Val	Tyr	Val	Ile	Pro	Asp	Gly	Glu	Thr	Arg
				185					190					195
Lys	Ile	Lys	His	Leu	Arg	Val	Phe	Phe	Ile	Thr	Ala	Ala	Trp	Ser
				200					205					210
Ile	Phe	Ala	Tyr	Ile	Trp	Leu	Tyr	Met	Ile	Leu	Ala	Val	Phe	Ser
				215					220					225
Pro	Gly	Val	Val	Gln	Val	Trp	Glu	Gly	Leu	Leu	Thr	Leu	Phe	Phe
				230					235					240
Phe	Pro	Val	Cys	Val	Leu	Leu	Ala	Trp	Val	Ala	Asp	Lys	Arg	Leu
				245					250					255
Leu	Phe	Tyr	Lys	Tyr	Met	His	Lys	Lys	Tyr	Arg	Thr	Asp	Lys	His
				260					265					270
Arg	Gly	Ile	Ile	Ile	Glu	Thr	Glu	Gly	Asp	His	Pro	Lys	Gly	Ile
				275					280					285
Glu	Met	Asp	Gly	Lys	Met	Met	Asn	Ser	His	Phe	Leu	Asp	Gly	Asn
				290					295					300
Leu	Val	Pro	Leu	Glu	Gly	Lys	Glu	Val	Asp	Glu	Ser	Arg	Arg	Glu
				305					310					315

Met	Ile	Arg	Ile	Leu	Lys	Asp	Leu	Lys	Gln	Lys	His	Pro	Glu	Lys
				320					325					330
Asp	Leu	Asp	Gln	Leu	Val	Glu	Met	Ala	Asn	Tyr	Tyr	Ala	Leu	Ser
				335					340					345
His	Gln	Gln	Lys	Ser	Arg	Ala	Phe	Tyr	Arg	Ile	Gln	Ala	Thr	Arg
				350					355					360
Met	Met	Thr	Gly	Ala	Gly	Asn	Ile	Leu	Lys	Lys	His	Ala	Ala	Glu
				365					370					375
Gln	Ala	Lys	Lys	Ala	Ser	Ser	Met	Ser	Glu	Val	His	Thr	Asp	Glu
				380					385					390
Pro	Glu	Asp	Phe	Ile	Ser	Lys	Val	Phe	Phe	Asp	Pro	Cys	Ser	Tyr
				395					400					405
Gln	Cys	Leu	Glu	Asn	Cys	Gly	Ala	Val	Leu	Leu	Thr	Val	Val	Arg
				410					415					420
Lys	Gly	Gly	Asp	Met	Ser	Lys	Thr	Met	Tyr	Val	Asp	Tyr	Lys	Thr
				425					430					435
Glu	Asp	Gly	Ser	Ala	Asn	Ala	Gly	Ala	Asp	Tyr	Glu	Phe	Thr	Glu
				440					445					450
Gly	Thr	Val	Val	Leu	Lys	Pro	Gly	Glu	Thr	Gln	Lys	Glu	Phe	Ser
				455					460					465
Val	Gly	Ile	Ile	Asp	Asp	Asp	Ile	Phe	Glu	Glu	Asp	Glu	His	Phe
				470					475					480
Phe	Val	Arg	Leu	Ser	Asn	Val	Arg	Ile	Glu	Glu	Glu	Gln	Pro	Glu
				485					490					495
Glu	Gly	Met	Pro	Pro	Ala	Ile	Phe	Asn	Ser	Leu	Pro	Leu	Pro	Arg
				500					505					510
Ala	Val	Leu	Ala	Ser	Pro	Cys	Val	Ala	Thr	Val	Thr	Ile	Leu	Asp
				515					520					525
Asp	Asp	His	Ala	Gly	Ile	Phe	Thr	Phe	Glu	Cys	Asp	Thr	Ile	His
				530					535					540
Val	Ser	Glu	Ser	Ile	Gly	Val	Met	Glu	Val	Lys	Val	Leu	Arg	Thr
				545					550					555
Ser	Gly	Ala	Arg	Gly	Thr	Val	Ile	Val	Pro	Phe	Arg	Thr	Val	Glu
				560					565					570
Gly	Thr	Ala	Lys	Gly	Gly	Gly	Glu	Asp	Phe	Glu	Asp	Thr	Tyr	Gly
				575					580					585
Glu	Leu	Glu	Phe	Lys	Asn	Asp	Glu	Thr	Val	Lys	Thr	Ile	Arg	Val
				590					595					600
Lys	Ile	Val	Asp	Glu	Glu	Glu	Tyr	Glu	Arg	Gln	Glu	Asn	Phe	Phe
				605					610					615
Ile	Ala	Leu	Gly	Glu	Pro	Lys	Trp	Met	Glu	Arg	Gly	Ile	Ser	Asp
				620					625					630
Val	Thr	Asp	Arg	Lys	Leu	Thr	Met	Glu	Glu	Glu	Glu	Ala	Lys	Arg
				635					640					645
Ile	Ala	Glu	Met	Gly	Lys	Pro	Val	Leu	Gly	Glu	His	Pro	Lys	Leu
				650					655					660
Glu	Val	Ile	Ile	Glu	Glu	Ser	Tyr	Glu	Phe	Lys	Thr	Thr	Val	Asp
				665					670					675
Lys	Leu	Ile	Lys	Lys	Thr	Asn	Leu	Ala	Leu	Val	Val	Gly	Thr	His
				680					685					690
Ser	Trp	Arg	Asp	Gln	Phe	Met	Glu	Ala	Ile	Thr	Val	Ser	Ala	Ala
				695					700					705
Gly	Asp	Glu	Asp	Glu	Asp	Glu	Ser	Gly	Glu	Glu	Arg	Leu	Pro	Ser
				710					715					720
Cys	Phe	Asp	Tyr	Val	Met	His	Phe	Leu	Thr	Val	Phe	Trp	Lys	Val
				725					730					735
Leu	Phe	Ala	Cys	Val	Pro	Pro	Thr	Glu	Tyr	Cys	His	Gly	Trp	Ala
				740					745					750
Cys	Phe	Ala	Val	Ser	Ile	Leu	Ile	Ile	Gly	Met	Leu	Thr	Ala	Ile
				755					760					765
Ile	Gly	Asp	Leu	Ala	Ser	His	Phe	Gly	Cys	Thr	Ile	Gly	Leu	Lys
				770					775					780
Asp	Ser	Val	Thr	Ala	Val	Val	Phe	Val	Ala	Phe	Gly	Thr	Ser	Val

Pro Asp Thr Phe	785	Ala Ser Lys Ala Ala	790	Ala Leu Gln Asp Val Tyr	795
	800		805		810
Ala Asp Ala Ser	815	Ile Gly Asn Val Thr	820	Gly Ser Asn Ala Val Asn	825
Val Phe Leu Gly	830	Ile Gly Leu Ala Trp	835	Ser Val Ala Ala Ile Tyr	840
Trp Ala Leu Gln	845	Gly Gln Glu Phe His	850	Ser Ala Gly Thr Leu	855
Ala Phe Ser Val	860	Thr Leu Phe Thr Ile	865	Phe Ala Phe Val Cys Ile	870
Ser Val Leu Leu	875	Tyr Arg Arg Arg Pro	880	His Leu Gly Gly Glu Leu	885
Gly Gly Pro Arg	890	Gly Cys Lys Leu Ala	895	Thr Thr Trp Leu Phe Val	900
Ser Leu Trp Leu	905	Leu Tyr Ile Leu Phe	910	Ala Thr Leu Glu Ala Tyr	915
Cys Tyr Ile Lys	920	Gly Phe			

<210> 10

<211> 1466

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6046484CD1

<400> 10

Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu	
1 5 10 15	
His Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu	
20 25 30	
Ile Ser Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys	
35 40 45	
Leu Lys Gln Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr	
50 55 60	
Thr Gly Ala Trp Ile Phe Thr Gly Gly Val Ser Thr Gly Val Ile	
65 70 75	
Ser His Val Gly Asp Ala Leu Lys Asp His Ser Ser Lys Ser Arg	
80 85 90	
Gly Arg Val Cys Ala Ile Gly Ile Ala Pro Trp Gly Ile Val Glu	
95 100 105	
Asn Lys Glu Asp Leu Val Gly Lys Asp Val Thr Arg Val Tyr Gln	
110 115 120	
Thr Met Ser Asn Pro Leu Ser Lys Leu Ser Val Leu Asn Asn Ser	
125 130 135	
His Thr His Phe Ile Leu Ala Asp Asn Gly Thr Leu Gly Lys Tyr	
140 145 150	
Gly Ala Glu Val Lys Leu Arg Arg Gln Leu Glu Lys His Ile Ser	
155 160 165	
Leu Gln Lys Ile Asn Thr Arg Ile Gly Gln Gly Val Pro Val Val	
170 175 180	
Ala Leu Ile Val Glu Gly Gly Pro Asn Val Ile Ser Ile Val Leu	
185 190 195	
Glu Tyr Leu Arg Asp Thr Pro Pro Val Pro Val Val Val Cys Asp	
200 205 210	
Gly Ser Gly Arg Ala Ser Asp Ile Leu Ala Phe Gly His Lys Tyr	
215 220 225	
Ser Glu Glu Gly Gly Leu Ile Asn Glu Ser Leu Arg Asp Gln Leu	
230 235 240	
Leu Val Thr Ile Gln Lys Thr Phe Thr Tyr Thr Arg Thr Gln Ala	



	245		250		255
Gln His Leu Phe	Ile Ile Leu Met Glu	Cys Met Lys Lys Lys	Glu		
	260		265		270
Leu Ile Thr Val	Phe Arg Met Gly Ser	Glu Gly His Gln Asp	Ile		
	275		280		285
Asp Leu Ala Ile	Leu Thr Ala Leu Leu	Lys Gly Ala Asn Ala	Ser		
	290		295		300
Ala Pro Asp Gln	Leu Ser Leu Ala Leu	Ala Trp Asn Arg Val	Asp		
	305		310		315
Ile Ala Arg Ser	Gln Ile Phe Ile Tyr	Gly Gln Gln Trp Pro	Val		
	320		325		330
Gly Ser Leu Glu	Gln Ala Met Leu Asp	Ala Leu Val Leu Asp	Arg		
	335		340		345
Val Asp Phe Val	Lys Leu Leu Ile Glu	Asn Gly Val Ser Met	His		
	350		355		360
Arg Phe Leu Thr	Ile Ser Arg Leu Glu	Glu Leu Tyr Asn Thr	Arg		
	365		370		375
His Gly Pro Ser	Asn Thr Leu Tyr His	Leu Val Arg Asp Val	Lys		
	380		385		390
Lys Gly Asn Leu	Pro Pro Asp Tyr Arg	Ile Ser Leu Ile Asp	Ile		
	395		400		405
Gly Leu Val Ile	Glu Tyr Leu Met Gly	Gly Ala Tyr Arg Cys	Asn		
	410		415		420
Tyr Thr Arg Lys	Arg Phe Arg Thr Leu	Tyr His Asn Leu Phe	Gly		
	425		430		435
Pro Lys Arg Asp	Asp Ile Pro Leu Arg	Arg Gly Arg Lys Thr	Thr		
	440		445		450
Lys Lys Arg Glu	Glu Glu Val Asp Ile	Asp Leu Asp Asp Pro	Glu		
	455		460		465
Ile Asn His Phe	Pro Phe Pro Phe His	Glu Leu Met Val Trp	Ala		
	470		475		480
Val Leu Met Lys	Arg Gln Lys Met Ala	Leu Phe Phe Trp Gln	His		
	485		490		495
Gly Glu Glu Ala	Met Ala Lys Ala Leu	Val Ala Cys Lys Leu	Cys		
	500		505		510
Lys Ala Met Ala	His Glu Ala Ser Glu	Asn Asp Met Val Asp	Asp		
	515		520		525
Ile Ser Gln Glu	Leu Asn His Asn Ser	Arg Asp Phe Gly Gln	Leu		
	530		535		540
Ala Val Glu Leu	Leu Asp Gln Ser Tyr	Lys Gln Asp Glu Gln	Leu		
	545		550		555
Ala Met Lys Leu	Leu Thr Tyr Glu Leu	Lys Asn Trp Ser Asn	Ala		
	560		565		570
Thr Cys Leu Gln	Leu Ala Val Ala Ala	Lys His Arg Asp Phe	Ile		
	575		580		585
Ala His Thr Cys	Ser Gln Met Leu Leu	Thr Asp Met Trp Met	Gly		
	590		595		600
Arg Leu Arg Met	Arg Lys Asn Ser Gly	Leu Lys Val Ile Leu	Gly		
	605		610		615
Ile Leu Leu Pro	Pro Thr Ile Leu Phe	Leu Glu Phe Arg Thr	Tyr		
	620		625		630
Asp Asp Phe Ser	Tyr Gln Thr Ser Lys	Glu Asn Glu Asp Gly	Lys		
	635		640		645
Glu Lys Glu Glu	Glu Asn Thr Asp Ala	Asn Ala Asp Ala Gly	Ser		
	650		655		660
Arg Lys Gly Asp	Glu Glu Asn Glu His	Lys Lys Gln Arg Ser	Ile		
	665		670		675
Pro Ile Gly Thr	Lys Ile Cys Glu Phe	Tyr Asn Ala Pro Ile	Val		
	680		685		690
Lys Phe Trp Phe	Phe Gln Ile Ser Tyr	Leu Gly Tyr Leu Leu	Leu		
	695		700		705
Phe Asn Tyr Val	Ile Leu Val Arg Met	Asp Gly Trp Pro Ser	Leu		
	710		715		720

Gln	Glu	Trp	Ile	Val	Ile	Ser	Tyr	Ile	Val	Ser	Leu	Ala	Leu	Glu	725	730	735
Lys	Ile	Arg	Glu	Val	Ala	Thr	Pro	Lys	Ala	Ser	Pro	Ser	Pro	Leu	740	745	750
Ala	Arg	Lys	Ser	Lys	Phe	Gly	Phe	Gln	Glu	Tyr	Trp	Asn	Ile	Thr	755	760	765
Asp	Leu	Val	Ala	Ile	Ser	Thr	Phe	Met	Ile	Gly	Ala	Ile	Leu	Arg	770	775	780
Leu	Gln	Asn	Gln	Pro	Tyr	Met	Gly	Tyr	Gly	Arg	Val	Ile	Tyr	Cys	785	790	795
Val	Asp	Ile	Ile	Phe	Trp	Tyr	Ile	Arg	Val	Leu	Asp	Ile	Phe	Gly	800	805	810
Val	Asn	Lys	Tyr	Leu	Gly	Pro	Tyr	Val	Met	Met	Ile	Gly	Lys	Met	815	820	825
Val	Ser	Ser	Gly	Ile	Leu	Trp	Val	Val	Ile	Met	Leu	Val	Val	Leu	830	835	840
Met	Ser	Phe	Gly	Val	Ala	Arg	Gln	Ala	Ile	Leu	His	Pro	Glu	Glu	845	850	855
Lys	Pro	Ser	Trp	Lys	Leu	Ala	Arg	Asn	Ile	Phe	Tyr	Met	Pro	Tyr	860	865	870
Trp	Met	Ile	Tyr	Gly	Glu	Val	Phe	Ala	Asp	Gln	Ile	Asp	Arg	Lys	875	880	885
Ser	Phe	Phe	Leu	Ser	Ala	Pro	Cys	Gly	Glu	Asn	Leu	Tyr	Asp	Glu	890	895	900
Glu	Gly	Lys	Arg	Leu	Pro	Pro	Cys	Ile	Pro	Gly	Ala	Trp	Leu	Thr	905	910	915
Pro	Ala	Leu	Met	Ala	Cys	Tyr	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	920	925	930
Val	Asn	Leu	Leu	Ile	Ala	Val	Phe	Ser	Asn	Thr	Phe	Phe	Glu	Val	935	940	945
Lys	Ser	Ile	Ser	Asn	Gln	Val	Trp	Lys	Phe	Gln	Arg	Tyr	Gln	Leu	950	955	960
Ile	Met	Thr	Phe	His	Asp	Arg	Pro	Val	Leu	Pro	Pro	Pro	Met	Ile	965	970	975
Ile	Leu	Ser	His	Ile	Tyr	Ile	Ile	Ile	Met	Arg	Leu	Ser	Gly	Arg	980	985	990
Cys	Arg	Lys	Lys	Arg	Glu	Gly	Asp	Gln	Glu	Glu	Arg	Asp	Arg	Gly	995	1000	1005
Leu	Ser	Met	Phe	Leu	Ser	Asp	Glu	Glu	Leu	Lys	Arg	Leu	His	Glu	1010	1015	1020
Phe	Glu	Glu	Gln	Cys	Val	Gln	Glu	His	Phe	Arg	Glu	Lys	Glu	Asp	1025	1030	1035
Glu	Gln	Gln	Ser	Ser	Ser	Asp	Glu	Arg	Ile	Arg	Val	Thr	Ser	Glu	1040	1045	1050
Arg	Val	Glu	Asn	Met	Ser	Met	Arg	Leu	Glu	Glu	Ile	Asn	Glu	Arg	1055	1060	1065
Glu	Thr	Phe	Met	Lys	Thr	Ser	Leu	Gln	Thr	Val	Asp	Leu	Arg	Leu	1070	1075	1080
Ala	Gln	Leu	Glu	Glu	Leu	Ser	Asn	Arg	Met	Val	Asn	Ala	Leu	Glu	1085	1090	1095
Asn	Leu	Ala	Gly	Ile	Asp	Arg	Ser	Asp	Leu	Ile	Gln	Ala	Arg	Ser	1100	1105	1110
Arg	Ala	Ser	Ser	Glu	Cys	Glu	Ala	Thr	Tyr	Leu	Leu	Arg	Gln	Ser	1115	1120	1125
Ser	Ile	Asn	Ser	Ala	Asp	Gly	Tyr	Ser	Leu	Tyr	Arg	Tyr	His	Phe	1130	1135	1140
Asn	Gly	Glu	Glu	Leu	Leu	Phe	Glu	Asp	Thr	Ser	Leu	Ser	Thr	Ser	1145	1150	1155
Pro	Gly	Thr	Gly	Val	Arg	Lys	Lys	Thr	Cys	Ser	Phe	Arg	Ile	Lys	1160	1165	1170
Glu	Glu	Lys	Asp	Val	Lys	Thr	His	Leu	Val	Pro	Glu	Cys	Gln	Asn	1175	1180	1185
Ser	Leu	His	Leu	Ser	Leu	Gly	Thr	Ser	Thr	Ser	Ala	Thr	Pro	Asp			

1190	1195	1200
Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu Glu Ser		
1205	1210	1215
Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu Arg		
1220	1225	1230
Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn		
1235	1240	1245
Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn		
1250	1255	1260
Thr Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser		
1265	1270	1275
Tyr Pro Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu		
1280	1285	1290
Thr Ile Asn Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr		
1295	1300	1305
Ser Arg Gly Arg Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu		
1310	1315	1320
Tyr Ser Ser Ile Thr Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys		
1325	1330	1335
Gln Val Gln Lys Ile Thr Arg Ser His Ser Thr Asp Ile Pro Tyr		
1340	1345	1350
Ile Val Ser Glu Ala Ala Val Gln Ala Glu His Lys Glu Gln Phe		
1355	1360	1365
Ala Asp Met Gln Asp Glu His His Val Ala Glu Ala Ile Pro Arg		
1370	1375	1380
Ile Pro Arg Leu Ser Leu Thr Ile Thr Asp Arg Asn Gly Met Glu		
1385	1390	1395
Asn Leu Leu Ser Val Lys Pro Asp Gln Thr Leu Gly Phe Pro Ser		
1400	1405	1410
Leu Arg Ser Lys Ser Leu His Gly His Pro Arg Asn Val Lys Ser		
1415	1420	1425
Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser Ser Val Ser		
1430	1435	1440
Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys Lys Val		
1445	1450	1455
Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys		
1460	1465	

<210> 11  
 <211> 222  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7481427CD1

<400> 11

Met Thr Glu Gln Ala Ile Ser Phe Ala Lys Asp Phe Leu Ala Gly	
1 5 10 15	
Gly Ile Thr Ala Ala Ile Ser Lys Thr Ala Val Ala Ser Ile Lys	
20 25 30	
Arg Val Gln Leu Leu Leu Gln Met Gln His Ala Ser Met Pro Met	
35 40 45	
Ala Ala Ala Lys Gln Cys Lys Gly Ile Val Asp Cys Ile Val Arg	
50 55 60	
Ile Pro Lys Asp Gln Gly Val Leu Ser Phe Trp Arg Gly Asn Leu	
65 70 75	
Ala Asn Val Ile Arg Tyr Ser Pro Thr Gln Ala Leu Asn Phe Ala	
80 85 90	
Phe Lys Asp Lys Tyr Lys Gln Ile Phe Leu Ala Gly Val Asp Lys	
95 100 105	
His Thr Gln Phe Cys Arg Tyr Phe Ala Gly Asn Leu Ala Ser Gly	

	110		115		120
Gly Thr Ala Val	Val Tyr Pro Leu Asp	Phe Thr Arg Thr Arg	Leu		
	125		130		135
Ala Ala Asp Val	Gly Lys Ser Gly Thr	Glu Arg Glu Phe Arg	Gly		
	140		145		150
Leu Gly Asp Cys	Leu Val Lys Ile Ser	Lys Ser Asp Gly Ile	Arg		
	155		160		165
Gly Leu Tyr Gln	Gly Phe Ser Val Ser	Val Gln Ala Ile Ile	Ile		
	170		175		180
Tyr Gln Ala Ala	Tyr Phe Arg Val Tyr	Asp Thr Ala Asn Gly	Met		
	185		190		195
Phe Pro Asp Pro	Lys Asn Thr His Ile	Leu Val Ser Trp Met	Thr		
	200		205		210
Ala Gln Thr Val	Thr Ala Val Ala Gly	Val Leu Ser			
	215		220		

<210> 12

<211> 461

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7483595CD1

<400> 12

Met Glu Thr Leu	Leu Phe Leu Glu Ser	Ala Ile Gly Ala Ile	Ala		
1	5	10	15		
Gly Leu Lys Thr	Phe Tyr Pro Thr	Leu Val Phe Ser	Ser Leu Gln		
	20	25	30		
Leu Ile Met Gly	Val Leu Gly Leu Gly	Phe Ile Ala Thr	Tyr Leu		
	35	40	45		
Pro Glu Ser Ala	Met Ser Ala Tyr	Leu Ala Val Ala	Leu His		
	50	55	60		
Ile Met Leu Ser	Gln Leu Thr Phe	Ile Phe Gly Ile	Met Ile Ser		
	65	70	75		
Phe His Ala Gly	Pro Ile Ser Phe	Phe Tyr Asp Ile	Ile Asn Tyr		
	80	85	90		
Cys Val Ala Leu	Pro Lys Ala Asn	Ser Thr Ser Ile	Leu Val Phe		
	95	100	105		
Leu Thr Val Val	Val Ala Leu Arg	Ile Asn Lys Cys	Ile Arg Ile		
	110	115	120		
Ser Phe Asn Gln	Tyr Pro Ile Glu	Phe Pro Met Glu	Leu Phe Leu		
	125	130	135		
Ile Ile Leu Gln	Ala Phe Ser Leu	Ser Leu Val Ser	Ser Phe Leu		
	140	145	150		
Leu Ile Phe Leu	Gly Lys Lys Ile	Ala Ser Leu His	Asn Tyr Ser		
	155	160	165		
Val Asn Ser Asn	Gln Asp Leu Ile	Ala Ile Gly Leu	Cys Asn Val		
	170	175	180		
Val Ser Ser Phe	Phe Arg Ser Cys	Val Phe Thr Gly	Ala Ile Ala		
	185	190	195		
Arg Thr Ile Ile	Gln Asp Lys Ser	Gly Gly Ser Thr	Thr Val Cys		
	200	205	210		
Ile Ser Gly Arg	Arg Arg Ala Lys	Ile Leu Leu Leu	Gly Gln Ile		
	215	220	225		
Pro Asn Thr Asn	Ile Tyr Arg Ser	Ile Asn Asp Tyr	Arg Glu Ile		
	230	235	240		
Ile Thr Ile Pro	Gly Val Lys Ile	Phe Gln Cys Cys	Ser Ser Ile		
	245	250	255		
Thr Phe Val Asn	Val Tyr Tyr Leu	Lys His Lys Leu	Leu Lys Glu		
	260	265	270		
Val Asp Met Val	Lys Val Pro Leu	Lys Glu Glu Glu	Ile Phe Ser		

Leu Phe Asn Ser	275	Ser Asp Thr Asn Leu	280	Gln Gly Gly Lys Ile	285
	290		295	Cys	300
Arg Cys Phe Cys	305	Asn Cys Asp Asp Leu	310	Glu Pro Leu Pro Arg	315
Leu Tyr Thr Glu	320	Arg Phe Glu Asn Lys	325	Leu Asp Pro Glu Ala	330
Ser Val Asn Leu	335	Ile His Cys Ser His	340	Phe Glu Ser Met Asn	345
Ser Gln Thr Ala	350	Ser Glu Asp Gln Val	355	Pro Tyr Thr Val Ser	360
Val Ser Gln Lys	365	Asn Gln Gly Gln Gln	370	Tyr Glu Glu Val Glu	375
Val Trp Leu Pro	380	Asn Asn Ser Ser Arg	385	Asn Ser Ser Pro Gly	390
Pro Asp Val Ala	395	Glu Ser Gln Gly Arg	400	Arg Ser Leu Ile Pro	405
Ser Asp Ala Ser	410	Leu Leu Pro Ser Val	415	His Thr Ile Ile Leu	420
Phe Ser Met Val	425	His Tyr Val Asp Ser	430	Arg Gly Leu Val Val	435
Arg Gln Val Ser	440	Thr Glu Glu Ala Leu	445	Ala Gly Ala Leu Ile	450
Leu Leu Pro Ser	455	Gln Pro His Pro Asp	460	Pro Asp	

<210> 13

<211> 502

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3788427CD1

<400> 13

Met Ser Ile Val	Arg	Leu Ser Val His	Ala	Lys Trp Ile Met	Gly
1	5		10		15
Lys Val Thr Gly	Thr	Lys Met Gln Lys	Thr	Ala Lys Val Arg	Val
	20		25		30
Ile Arg Leu Val	Leu Asp Pro His	Leu Leu Lys	Tyr Tyr Asn	Lys	
	35		40		45
Gln Lys Thr Tyr	Phe Ala His Asn	Ala Leu Gln Gln	Cys Thr	Ile	
	50		55		60
Gly Asp Ile Val	Leu Leu Lys Ala	Leu Pro Val Pro	Arg Thr	Lys	
	65		70		75
His Val Lys His	Glu Leu Ala Glu	Ile Val Phe Lys	Val Gly	Lys	
	80		85		90
Leu Val Asp Pro	Val Thr Gly Lys	Pro Arg Ala Gly	Thr Thr	Tyr	
	95		100		105
Leu Glu Ser Pro	Leu Ser Ser Glu	Thr Thr Gln Gly	Val Asp	Gly	
	110		115		120
Ala Ser Arg Pro	Ser Arg Gly Pro	Ala Pro Cys Arg	Ala Gly	Pro	
	125		130		135
Gly Ala Arg Arg	Leu Arg Pro Trp	Pro Glu Ser Pro	Arg Pro	Glu	
	140		145		150
Pro Arg Gly Leu	Pro Gly Pro Gly	Arg Gly Ser Met	Ala Thr	Trp	
	155		160		165
Arg Arg Asp Gly	Arg Leu Thr Gly	Gly Gln Arg Leu	Leu Cys	Ala	
	170		175		180
Gly Leu Ala Gly	Thr Leu Ser Leu	Ser Leu Thr Ala	Pro Leu	Glu	
	185		190		195
Leu Ala Thr Val	Leu Ala Gln Val	Gly Val Val Arg	Gly His	Ala	

Arg Gly Pro Trp	200	205	210
Ala Thr Gly His Arg Val Trp Arg Ala Glu Gly	215	220	225
Leu Arg Ala Leu Trp Lys Gly Asn Ala Val Ala Cys Leu Arg Leu	230	235	240
Phe Pro Cys Ser Ala Val Gln Leu Ala Ala Tyr Arg Lys Phe Val	245	250	255
Val Leu Phe Thr Asp Asp Leu Gly His Ile Ser Gln Trp Ser Ser	260	265	270
Ile Met Ala Gly Ser Leu Ala Gly Met Val Ser Thr Ile Val Thr	275	280	285
Tyr Pro Thr Asp Leu Ile Lys Thr Arg Leu Ile Met Gln Asn Ile	290	295	300
Leu Glu Pro Ser Tyr Arg Gly Leu Leu His Ala Phe Ser Thr Ile	305	310	315
Tyr Gln Gln Glu Gly Phe Leu Ala Leu Tyr Arg Gly Val Ser Leu	320	325	330
Thr Val Val Gly Ala Leu Pro Phe Ser Ala Gly Ser Leu Leu Val	335	340	345
Tyr Met Asn Leu Glu Lys Ile Trp Asn Gly Pro Arg Asp Gln Phe	350	355	360
Ser Leu Pro Gln Asn Phe Ala Asn Val Cys Leu Ala Ala Ala Val	365	370	375
Thr Gln Thr Leu Ser Phe Pro Phe Glu Thr Val Lys Arg Lys Met	380	385	390
Gln Ala Gln Ser Pro Tyr Leu Pro His Ser Gly Gly Val Asp Val	395	400	405
His Phe Ser Gly Ala Val Asp Cys Phe Arg Gln Ile Val Lys Ala	410	415	420
Gln Gly Val Leu Gly Leu Trp Asn Gly Leu Thr Ala Asn Leu Leu	425	430	435
Lys Ile Val Pro Tyr Phe Gly Ile Met Phe Ser Thr Phe Glu Phe	440	445	450
Cys Lys Arg Ile Cys Leu Tyr Gln Asn Gly Tyr Ile Leu Ser Pro	455	460	465
Leu Ser Tyr Lys Leu Thr Pro Gly Val Asp Gln Ser Leu Gln Pro	470	475	480
Gln Glu Leu Arg Glu Leu Lys Lys Phe Phe Lys Thr Arg Lys Leu	485	490	495
Lys Pro Lys Lys Pro Thr Leu	500		

<210> 14

<211> 261

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6972455CD1

<400> 14

Met Thr Lys Arg Tyr Ser Ala Leu Leu Thr Ala Leu Phe Ala Ser	1	5	10	15
Leu Met Leu Ser Gln Ala Pro Ala Gln Ala Ser Gly Leu Asp Asp	20	25	30	35
Ile Val Ala Arg Gly Thr Leu Lys Val Ala Val Pro Gln Asp Phe	40	45	50	55
Pro Pro Phe Gly Ser Val Gly Pro Asp Met Gln Pro Arg Gly Leu	60	65	70	75
Asp Ile Asp Thr Ala Lys Leu Leu Ala Asp Gln Leu Lys Val Lys	80	85	90	95
Leu Glu Leu Thr Pro Val Asn Ser Thr Asn Arg Ile Pro Phe Leu	100	105	110	115

	80		85		90
Thr Thr Gly Lys Val Asp Leu Val Ile Ser Ser Leu Gly Lys Asn					
	95		100		105
Pro Glu Arg Ala Lys Val Ile Asp Phe Ser Asn Ala Tyr Ala Pro					
	110		115		120
Phe Tyr Leu Ala Val Phe Gly Pro Pro Asp Ala Ala Ile Ala Ser					
	125		130		135
Leu Asp Asp Leu Lys Gly Lys Thr Ile Ser Val Thr Arg Gly Ala					
	140		145		150
Ile Glu Asp Ile Glu Leu Thr Ala Val Ala Pro Lys Glu Ala Thr					
	155		160		165
Ile Lys Arg Phe Glu Asp Asn Asn Ser Thr Ile Ala Ala Tyr Leu					
	170		175		180
Ala Gly Gln Val Asp Leu Ile Ala Ser Gly Asn Val Val Met Val					
	185		190		195
Ala Ile Ser Glu Arg Asn Pro Lys Arg Val Pro Ala Leu Lys Val					
	200		205		210
Lys Leu Lys Asp Ser Pro Val Tyr Val Gly Val Asn Lys Asn Glu					
	215		220		225
Pro Ala Leu Leu Glu Lys Val Asn Gln Ile Leu Val Ala Ala Lys					
	230		235		240
Ala Asp Gly Ser Leu Gly Lys Asn Ala Met Gln Trp Leu Lys Glu					
	245		250		255
Pro Leu Pro Ala Asp Leu					
	260				

<210> 15  
 <211> 570  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 8077668CD1

<400> 15	
Met Tyr Asp Pro Pro Lys Val Asn Lys Lys Asn Asp Ser Ala Asp	
1 5 10 15	
His Ser Asp Glu Ser Val Glu Ile Gln Ser Ala Leu His Ile Cys	
20 25 30	
Gly Phe Asp Gln Gly Arg Trp Ile Leu Ser Tyr Arg Arg Ser Pro	
35 40 45	
Ala Arg Arg His Thr Gly Leu Ala Pro Val Leu Gly Val Arg Val	
50 55 60	
Arg Val Cys Val Gly Cys Gln Gly Ala Arg Val Pro Ala Trp Pro	
65 70 75	
Lys Arg Glu Gly Arg Gly Ser Ala Glu Glu Pro Val Glu Arg Lys	
80 85 90	
Asp Gly Gln Gln Gly Glu Arg Arg Gly Glu Ala Ser Leu Thr Lys	
95 100 105	
Asp Ala Val Thr Phe Gln Ser Ser Leu Pro Val Gln Ile Thr Arg	
110 115 120	
Ile Pro Ala Cys Trp Arg Met Thr Gln Met Ser Gln Val Gln Glu	
125 130 135	
Leu Phe His Glu Ala Ala Gln Gln Asp Ala Leu Ala Gln Pro Gln	
140 145 150	
Pro Trp Trp Lys Thr Gln Leu Phe Met Trp Glu Pro Val Leu Phe	
155 160 165	
Gly Thr Trp Asp Gly Val Phe Thr Ser Cys Met Ile Asn Ile Phe	
170 175 180	
Gly Val Val Leu Phe Leu Arg Thr Gly Trp Leu Val Gly Asn Thr	
185 190 195	
Gly Val Leu Leu Gly Met Phe Leu Val Ser Phe Val Ile Leu Val	

Ala Leu Val Thr	200	205	210
Val Leu Ser Asp Ile	215	Gly Val Gly Glu Arg Ser	225
Ser Ile Gly Ser	230	235	240
Gly Gly Gln Thr	245	250	255
Gln Met Tyr Ile	260	265	270
Gly Leu Gly Asn	275	280	285
Leu Leu Ala Leu	290	295	300
Ile Arg Leu Gln	305	310	315
Leu Asp Phe Val	320	325	330
Gly Phe Ile Gly	335	340	345
Pro Asp Tyr Ser	350	355	360
Phe Phe Pro Ala	365	370	375
Gly Asp Leu Arg	380	385	390
Ala Ala Val Gly	395	400	405
His Gly Arg Leu	410	415	420
Gln His Ser Pro	425	430	435
Ser Pro His Gly	440	445	450
Pro Arg Gln Lys	455	460	465
Thr Val Thr Asn	470	475	480
Gly Gly Leu Glu	485	490	495
Ser Cys Ala Leu	500	505	510
Pro Asp His Arg	515	520	525
Pro Pro Ser Leu	530	535	540
Asn His Arg Val	545	550	555
Ser Phe Ala Ile	560	565	570

<210> 16

<211> 1033

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 55120485CD1

<400> 16

Met Glu Glu Asn Ser	Lys Lys Asp His	Arg Ala Leu Leu Asn Gln
1	5	10 15



Gly	Glu	Glu	Asp	Glu	Leu	Glu	Val	Phe	Gly	Tyr	Arg	Asp	His	Asn	
				20					25					30	
Val	Arg	Lys	Ala	Phe	Cys	Leu	Val	Ala	Ser	Val	Leu	Thr	Cys	Gly	
				35					40					45	
Gly	Leu	Leu	Leu	Val	Phe	Tyr	Trp	Arg	Pro	Gln	Trp	Arg	Val	Trp	
				50					55					60	
Ala	Asn	Cys	Ile	Pro	Cys	Pro	Leu	Gln	Glu	Ala	Asp	Thr	Val	Leu	
				65					70					75	
Leu	Arg	Thr	Thr	Asp	Glu	Phe	Gln	Arg	Tyr	Met	Arg	Lys	Lys	Val	
				80					85					90	
Phe	Cys	Leu	Tyr	Leu	Tyr	Thr	Leu	Lys	Phe	Pro	Val	Ser	Lys	Lys	
				95					100					105	
Trp	Glu	Glu	Ser	Leu	Val	Ala	Asp	Arg	His	Ser	Val	Ile	Asn	Gln	
				110					115					120	
Ala	Leu	Ile	Lys	Pro	Glu	Leu	Lys	Leu	Arg	Cys	Leu	Glu	Val	Gln	
				125					130					135	
Lys	Ile	Arg	Tyr	Val	Trp	Asn	Asp	Leu	Glu	Lys	Arg	Phe	Gln	Lys	
				140					145					150	
Val	Gly	Leu	Leu	Glu	Asp	Ser	Asn	Ser	Cys	Ser	Asp	Ile	His	Gln	
				155					160					165	
Thr	Phe	Gly	Leu	Gly	Leu	Thr	Ser	Glu	Glu	Gln	Glu	Val	Arg	Arg	
				170					175					180	
Leu	Val	Cys	Gly	Pro	Asn	Ala	Ile	Glu	Val	Glu	Ile	Gln	Pro	Ile	
				185					190					195	
Trp	Lys	Leu	Leu	Val	Lys	Gln	Val	Leu	Asn	Pro	Phe	Tyr	Val	Phe	
				200					205					210	
Gln	Ala	Phe	Thr	Leu	Thr	Leu	Trp	Leu	Ser	Gln	Gly	Tyr	Ile	Glu	
				215					220					225	
Tyr	Ser	Val	Ala	Ile	Ile	Ile	Leu	Thr	Val	Ile	Ser	Ile	Val	Leu	
				230					235					240	
Ser	Val	Tyr	Asp	Leu	Arg	Gln	Gln	Ser	Val	Lys	Leu	His	Asn	Leu	
				245					250					255	
Val	Glu	Asp	His	Asn	Lys	Val	Gln	Val	Thr	Ile	Ile	Val	Lys	Asp	
				260					265					270	
Lys	Gly	Leu	Glu	Glu	Leu	Glu	Ser	Arg	Leu	Leu	Val	Pro	Gly	Asp	
				275					280					285	
Ile	Leu	Ile	Leu	Pro	Gly	Lys	Phe	Ser	Leu	Pro	Cys	Asp	Ala	Val	
				290					295					300	
Leu	Ile	Asp	Gly	Ser	Cys	Val	Val	Asn	Glu	Gly	Met	Leu	Thr	Gly	
				305					310					315	
Glu	Ser	Ile	Pro	Val	Thr	Lys	Thr	Pro	Leu	Pro	Gln	Met	Glu	Asn	
				320					325					330	
Thr	Met	Pro	Trp	Lys	Cys	His	Ser	Leu	Glu	Asp	Tyr	Arg	Lys	His	
				335					340					345	
Val	Leu	Phe	Cys	Gly	Thr	Glu	Val	Ile	Gln	Val	Lys	Pro	Ser	Gly	
				350					355					360	
Gln	Gly	Pro	Val	Arg	Ala	Val	Val	Leu	Gln	Thr	Gly	Tyr	Asn	Thr	
				365					370					375	
Ala	Lys	Gly	Asp	Leu	Val	Arg	Ser	Ile	Leu	Tyr	Pro	Arg	Pro	Leu	
				380					385					390	
Asn	Phe	Lys	Leu	Tyr	Ser	Asp	Ala	Phe	Lys	Phe	Ile	Val	Phe	Leu	
				395					400					405	
Ala	Cys	Leu	Gly	Val	Met	Gly	Phe	Phe	Tyr	Ala	Leu	Gly	Val	Tyr	
				410					415					420	
Met	Tyr	His	Gly	Val	Pro	Pro	Lys	Asp	Thr	Val	Thr	Met	Ala	Leu	
				425					430					435	
Ile	Leu	Leu	Thr	Val	Thr	Val	Pro	Pro	Val	Leu	Pro	Ala	Ala	Leu	
				440					445					450	
Thr	Ile	Gly	Asn	Val	Tyr	Ala	Gln	Lys	Arg	Leu	Lys	Lys	Lys	Lys	
				455					460					465	
Ile	Phe	Cys	Ile	Ser	Pro	Gln	Arg	Ile	Asn	Met	Cys	Gly	Gln	Ile	
				470					475					480	
Asn	Leu	Val	Cys	Phe	Asp	Lys	Thr	Gly	Thr	Leu	Thr	Glu	Asp	Gly	

Leu Asp Leu Trp	485	Gly Thr Val Pro Thr	490	Ala Asp Asn Cys Phe	495
500		505		510	
Glu Ala His Ser	515	Phe Ala Ser Gly Gln	520	Ala Val Pro Trp Ser	525
Leu Cys Ala Ala	530	Met Ala Ser Cys His	535	Ser Leu Ile Leu Leu	540
Gly Thr Ile Gln	545	Gly Asp Pro Leu Asp	550	Leu Lys Met Phe Glu	555
Thr Ala Trp Lys	560	Met Glu Asp Cys Ile	565	Val Asp Ser Cys Lys	570
Gly Thr Ser Val	575	Ser Asn Ile Ile Lys	580	Pro Gly Pro Lys Ala	585
Lys Ser Pro Val	590	Glu Ala Ile Ile Thr	595	Leu Cys Gln Phe Pro	600
Ser Ser Ser Leu	605	Gln Arg Met Ser Val	610	Ile Ala Gln Leu Ala	615
Glu Asn His Phe	620	His Val Tyr Met Lys	625	Gly Ala Pro Glu Met	630
Ala Arg Phe Cys	635	Arg Ser Glu Thr Val	640	Pro Lys Asn Phe Pro	645
Glu Leu Arg Ser	650	Tyr Thr Val Gln Gly	655	Phe Arg Val Ile Ala	660
Ala His Lys Thr	665	Leu Lys Met Gly Asn	670	Leu Ser Glu Val Glu	675
Leu Ala Arg Glu	680	Lys Val Glu Ser Glu	685	Leu Thr Phe Leu Gly	690
Leu Ile Met Glu	695	Asn Arg Leu Lys Lys	700	Glu Thr Lys Leu Val	705
Lys Glu Leu Ser	710	Glu Ala Arg Ile Arg	715	Thr Val Met Ile Thr	720
Asp Asn Leu Gln	725	Thr Ala Ile Thr Val	730	Ala Lys Asn Ser Glu	735
Ile Pro Pro Gly	740	Ser Gln Val Ile Ile	745	Val Glu Ala Asp Glu	750
Glu Glu Phe Val	755	Pro Ala Ser Val Thr	760	Trp Gln Leu Val Glu	765
Gln Glu Thr Gly	770	Pro Gly Lys Lys Glu	775	Ile Tyr Met His Thr	780
Asn Ser Ser Thr	785	Pro Arg Gly Glu Gly	790	Gly Ser Cys Tyr His	795
Ala Met Ser Gly	800	Lys Ser Tyr Gln Val	805	Ile Phe Gln His Phe	810
Ser Leu Leu Pro	815	Lys Ile Leu Val Asn	820	Gly Thr Val Phe Ala	825
Met Ser Pro Gly	830	Gln Lys Ser Ser Leu	835	Ile Glu Glu Phe Gln	840
Leu Asn Ala Cys	845	Thr Val Gln Asn Glu	850	Ser Ile Ser Glu Leu	855
Met Ser Pro Thr	860	Ala Pro Glu Lys Met	865	Glu Ser Asn Ser Thr	870
Thr Ser Phe Glu	875	Asn Thr Thr Val Trp	880	Phe Leu Gly Thr Ile	885
Cys Ile Thr Val	890	Ala Leu Val Phe Ser	895	Lys Gly Lys Pro Phe	900
Gln Pro Thr Tyr	905	Thr Asn Tyr Ile Phe	910	Val Leu Val Leu Ile	915
Gln Leu Gly Val	920	Cys Leu Phe Ile Leu	925	Phe Ala Asp Ile Pro	930
Leu Tyr Arg Arg	935	Leu Asp Leu Leu Cys	940	Thr Pro Val Leu Trp	945
Ala Ser Ile Val	950	Ile Met Leu Ser Leu	955	Asn Phe Ile Val Ser	960

Val	Ala	Glu	Glu	Ala	Val	Ile	Glu	Asn	Arg	Ala	Leu	Trp	Met	Met	
				965					970					975	
Ile	Lys	Arg	Cys	Phe	Gly	Tyr	Gln	Ser	Lys	Ser	Gln	Tyr	Arg	Ile	
				980					985					990	
Trp	Gln	Arg	Asp	Leu	Ala	Asn	Asp	Pro	Ser	Trp	Pro	Pro	Leu	Asn	
				995					1000					1005	
Gln	Thr	Ser	His	Ser	Asp	Met	Pro	Glu	Cys	Gly	Arg	Gly	Val	Ser	
				1010					1015					1020	
Tyr	Ser	Asn	Pro	Val	Phe	Glu	Ser	Asn	Glu	Glu	Gln	Leu			
				1025					1030						

<210> 17

<211> 496

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3112883CD1

<400> 17

Met	Asp	Cys	Tyr	Arg	Thr	Ser	Leu	Ser	Ser	Ser	Trp	Ile	Tyr	Pro	
1				5					10					15	
Thr	Val	Ile	Leu	Cys	Leu	Phe	Gly	Phe	Ser	Met	Met	Arg	Pro		
				20					25					30	
Ser	Glu	Pro	Phe	Leu	Ile	Pro	Tyr	Leu	Ser	Gly	Pro	Asp	Lys	Asn	
				35					40					45	
Leu	Thr	Ser	Ala	Glu	Ile	Thr	Asn	Glu	Ile	Phe	Pro	Val	Trp	Thr	
				50					55					60	
Tyr	Ser	Tyr	Leu	Val	Leu	Leu	Leu	Pro	Val	Phe	Val	Leu	Thr	Asp	
				65					70					75	
Tyr	Val	Arg	Tyr	Lys	Pro	Val	Ile	Ile	Leu	Gln	Gly	Ile	Ser	Phe	
				80					85					90	
Ile	Ile	Thr	Trp	Leu	Leu	Leu	Leu	Phe	Gly	Gln	Gly	Val	Lys	Thr	
				95					100					105	
Met	Gln	Val	Val	Glu	Phe	Phe	Tyr	Gly	Met	Val	Thr	Ala	Ala	Glu	
				110					115					120	
Val	Ala	Tyr	Tyr	Ala	Tyr	Ile	Tyr	Ser	Val	Val	Ser	Pro	Glu	His	
				125					130					135	
Tyr	Gln	Arg	Val	Ser	Gly	Tyr	Cys	Arg	Ser	Val	Thr	Leu	Ala	Ala	
				140					145					150	
Tyr	Thr	Ala	Gly	Ser	Val	Leu	Ala	Gln	Leu	Leu	Val	Ser	Leu	Ala	
				155					160					165	
Asn	Met	Ser	Tyr	Phe	Tyr	Leu	Asn	Val	Ile	Ser	Leu	Ala	Ser	Val	
				170					175					180	
Ser	Val	Ala	Phe	Leu	Phe	Ser	Leu	Phe	Leu	Pro	Met	Pro	Lys	Lys	
				185					190					195	
Ser	Met	Phe	Phe	His	Ala	Lys	Pro	Ser	Arg	Glu	Ile	Lys	Lys	Ser	
				200					205					210	
Ser	Ser	Val	Asn	Pro	Val	Leu	Glu	Glu	Thr	His	Glu	Gly	Glu	Ala	
				215					220					225	
Pro	Gly	Cys	Glu	Glu	Gln	Lys	Pro	Thr	Ser	Glu	Ile	Leu	Ser	Thr	
				230					235					240	
Ser	Gly	Lys	Leu	Asn	Lys	Gly	Gln	Leu	Asn	Ser	Leu	Lys	Pro	Ser	
				245					250					255	
Asn	Val	Thr	Val	Asp	Val	Phe	Val	Gln	Trp	Phe	Gln	Asp	Leu	Lys	
				260					265					270	
Glu	Cys	Tyr	Ser	Ser	Lys	Arg	Leu	Phe	Tyr	Trp	Ser	Leu	Trp	Trp	
				275					280					285	
Ala	Phe	Ala	Thr	Ala	Gly	Phe	Asn	Gln	Val	Leu	Asn	Tyr	Val	Gln	
				290					295					300	
Ile	Leu	Trp	Asp	Tyr	Lys	Ala	Pro	Ser	Gln	Asp	Ser	Ser	Ile	Tyr	
				305					310					315	

Asn	Gly	Ala	Val	Glu	Ala	Ile	Ala	Thr	Phe	Gly	Gly	Ala	Val	Ala
				320					325					330
Ala	Phe	Ala	Val	Gly	Tyr	Val	Lys	Val	Asn	Trp	Asp	Leu	Leu	Gly
				335					340					345
Glu	Leu	Ala	Leu	Val	Val	Phe	Ser	Val	Val	Asn	Ala	Gly	Ser	Leu
				350					355					360
Phe	Leu	Met	His	Tyr	Thr	Ala	Asn	Ile	Trp	Ala	Cys	Tyr	Ala	Gly
				365					370					375
Tyr	Leu	Ile	Phe	Lys	Ser	Ser	Tyr	Met	Leu	Leu	Ile	Thr	Ile	Ala
				380					385					390
Val	Phe	Gln	Ile	Ala	Val	Asn	Leu	Asn	Val	Glu	Arg	Tyr	Ala	Leu
				395					400					405
Val	Phe	Gly	Ile	Asn	Thr	Phe	Ile	Ala	Leu	Val	Ile	Gln	Thr	Ile
				410					415					420
Met	Thr	Val	Ile	Val	Val	Asp	Gln	Arg	Gly	Leu	Asn	Leu	Pro	Val
				425					430					435
Ser	Ile	Gln	Phe	Leu	Val	Tyr	Gly	Ser	Tyr	Phe	Ala	Val	Ile	Ala
				440					445					450
Gly	Ile	Phe	Leu	Met	Arg	Ser	Met	Tyr	Ile	Thr	Tyr	Ser	Thr	Lys
				455					460					465
Ser	Gln	Lys	Asp	Val	Gln	Ser	Pro	Ala	Pro	Ser	Glu	Asn	Pro	Asp
				470					475					480
Val	Ser	His	Pro	Glu	Glu	Glu	Ser	Asn	Ile	Ile	Met	Ser	Thr	Lys
				485					490					495

Leu

<210> 18  
 <211> 573  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4253888CD1

<400> 18

Met	Thr	Asn	Ser	Lys	Gly	Arg	Ser	Ile	Thr	Asp	Lys	Thr	Ser	Gly
1				5					10					15
Gly	Pro	Ser	Ser	Gly	Gly	Gly	Phe	Val	Asp	Trp	Thr	Leu	Arg	Leu
				20					25					30
Asn	Thr	Ile	Gln	Ser	Asp	Lys	Phe	Leu	Asn	Leu	Leu	Leu	Ser	Met
				35					40					45
Val	Pro	Val	Ile	Tyr	Gln	Lys	Asn	Gln	Glu	Asp	Arg	His	Lys	Lys
				50					55					60
Pro	Asn	Gly	Ile	Trp	Gln	Asp	Gly	Leu	Ser	Thr	Ala	Val	Gln	Thr
				65					70					75
Phe	Ser	Asn	Arg	Ser	Glu	Gln	His	Met	Glu	Tyr	His	Ser	Phe	Ser
				80					85					90
Glu	Gln	Ser	Phe	His	Ala	Asn	Asn	Gly	His	Ala	Ser	Ser	Ser	Cys
				95					100					105
Ser	Gln	Lys	Tyr	Asp	Asp	Tyr	Ala	Asn	Tyr	Asn	Tyr	Cys	Asp	Gly
				110					115					120
Arg	Glu	Thr	Ser	Glu	Thr	Thr	Ala	Met	Leu	Gln	Asp	Glu	Asp	Ile
				125					130					135
Ser	Ser	Asp	Gly	Asp	Glu	Asp	Ala	Ile	Val	Glu	Val	Thr	Pro	Lys
				140					145					150
Leu	Pro	Lys	Glu	Ser	Ser	Gly	Ile	Met	Ala	Leu	Gln	Ile	Leu	Val
				155					160					165
Pro	Phe	Leu	Leu	Ala	Gly	Phe	Gly	Thr	Val	Ser	Ala	Gly	Met	Val
				170					175					180
Leu	Asp	Ile	Val	Gln	His	Trp	Glu	Val	Phe	Arg	Lys	Val	Thr	Glu
				185					190					195

Val Phe Ile Leu	Val Pro Ala Leu Leu	Gly Leu Lys Gly Asn	Leu
200	205	210	
Glu Met Thr Leu	Ala Ser Arg Leu Ser	Thr Ala Val Asn Ile	Gly
215	220	225	
Lys Met Asp Ser	Pro Ile Glu Lys Trp	Asn Leu Ile Ile Gly	Asn
230	235	240	
Leu Ala Leu Lys	Gln Val Gln Ala Thr	Val Val Gly Phe Leu	Ala
245	250	255	
Ala Val Ala Ala	Ile Ile Leu Gly Trp	Ile Pro Glu Gly Lys	Tyr
260	265	270	
Tyr Leu Asp His	Ser Ile Leu Leu Cys	Ser Ser Ser Val Ala	Thr
275	280	285	
Ala Phe Ile Ala	Ser Leu Leu Gln Gly	Ile Ile Met Val Gly	Val
290	295	300	
Ile Val Gly Ser	Lys Lys Thr Gly Ile	Asn Pro Asp Asn Val	Ala
305	310	315	
Thr Pro Ile Ala	Ala Ser Phe Gly Asp	Leu Ile Thr Leu Ala	Ile
320	325	330	
Leu Ala Trp Ile	Ser Gln Gly Leu Tyr	Ser Cys Leu Glu Thr	Tyr
335	340	345	
Tyr Tyr Ile Ser	Pro Leu Val Gly Val	Phe Phe Leu Ala Leu	Thr
350	355	360	
Pro Ile Trp Ile	Ile Ile Ala Ala Lys	His Pro Ala Thr Arg	Thr
365	370	375	
Val Leu His Ser	Gly Trp Glu Pro Val	Ile Thr Ala Met Val	Ile
380	385	390	
Ser Ser Ile Gly	Gly Leu Ile Leu Asp	Thr Thr Val Ser Asp	Pro
395	400	405	
Asn Leu Val Gly	Ile Val Val Tyr Thr	Pro Val Ile Asn Gly	Ile
410	415	420	
Gly Gly Asn Leu	Val Ala Ile Gln Ala	Ser Arg Ile Ser Thr	Tyr
425	430	435	
Leu His Leu His	Ser Ile Pro Gly Glu	Leu Pro Asp Glu Pro	Lys
440	445	450	
Gly Cys Tyr Tyr	Pro Phe Arg Thr Phe	Phe Gly Pro Gly Val	Asn
455	460	465	
Asn Lys Ser Ala	Gln Val Leu Leu Leu	Leu Val Ile Pro Gly	His
470	475	480	
Leu Ile Phe Leu	Tyr Thr Ile His Leu	Met Lys Ser Gly His	Thr
485	490	495	
Ser Leu Thr Ile	Ile Phe Ile Val Val	Tyr Leu Phe Gly Ala	Val
500	505	510	
Leu Gln Val Phe	Thr Leu Leu Trp Ile	Ala Asp Trp Met Val	His
515	520	525	
His Phe Trp Arg	Lys Gly Lys Asp Pro	Asp Ser Phe Ser Ile	Pro
530	535	540	
Tyr Leu Thr Ala	Leu Gly Asp Leu Leu	Gly Thr Ala Leu Leu	Ala
545	550	555	
Leu Ser Phe His	Phe Leu Trp Leu Ile	Gly Asp Arg Asp Gly	Asp
560	565	570	
Val Gly Asp			

<210> 19  
 <211> 573  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7479974CD1

<400> 19

Met	Asp	Ala	Val	Lys	Tyr	Leu	Asn	Lys	Leu	Asn	Leu	Asp	Asn	Ile
1				5					10					15
Glu	Leu	Thr	Lys	Tyr	Leu	Phe	Phe	Thr	Gly	Lys	Gly	Gly	Val	Gly
				20					25					30
Lys	Thr	Thr	Ile	Ser	Ser	Phe	Ile	Ala	Leu	Asn	Leu	Ala	Glu	Asn
				35					40					45
Gly	Lys	Lys	Val	Ala	Leu	Val	Ser	Thr	Asp	Pro	Ala	Ser	Asn	Leu
				50					55					60
Gln	Asp	Val	Phe	Gln	Met	Glu	Leu	Ser	Asn	Lys	Leu	Thr	Lys	Tyr
				65					70					75
Gln	Pro	Ile	Pro	Asn	Leu	Ser	Ile	Ala	Asn	Phe	Asp	Pro	Ile	Val
				80					85					90
Ala	Ala	Asp	Asp	Tyr	Lys	Ala	Gln	Ser	Ile	Glu	Pro	Tyr	Glu	Gly
				95					100					105
Ile	Leu	Pro	Glu	Asp	Val	Leu	Ala	Glu	Met	Lys	Glu	Gln	Leu	Ser
				110					115					120
Gly	Ser	Cys	Thr	Val	Glu	Val	Ala	Ala	Phe	Asn	Glu	Phe	Thr	Asn
				125					130					135
Phe	Leu	Ser	Asp	Lys	Thr	Leu	Glu	Gln	Glu	Phe	Asp	Phe	Ile	Ile
				140					145					150
Phe	Asp	Thr	Ala	Pro	Thr	Gly	His	Thr	Leu	Arg	Met	Leu	Glu	Leu
				155					160					165
Pro	Ser	Ala	Trp	Thr	Asp	Tyr	Leu	Asn	Thr	Thr	Ser	Asn	Asp	Ala
				170					175					180
Ser	Cys	Leu	Gly	Gln	Leu	Ser	Gly	Leu	Asn	Glu	Asn	Arg	Val	Lys
				185					190					195
Tyr	Asn	Ser	Ala	Leu	Glu	Lys	Leu	Arg	Asn	Gln	Asp	Asp	Thr	Thr
				200					205					210
Met	Met	Leu	Val	Ala	Arg	Pro	Thr	His	Ser	Ser	Ile	Tyr	Glu	Ile
				215					220					225
Gln	Arg	Ala	Gln	Gln	Glu	Leu	Gln	Gln	Leu	Ser	Ile	Ser	Lys	Phe
				230					235					240
Lys	Val	Ile	Ile	Asn	Asn	Tyr	Ile	Glu	Glu	Ser	His	Gly	Leu	Ile
				245					250					255
Ser	Ser	Gln	Met	Lys	Ser	Glu	Gln	Asp	Lys	Asn	Ile	Asn	His	Phe
				260					265					270
Thr	Glu	Trp	Leu	Asn	Asn	Asn	His	Ala	Tyr	Tyr	Val	Pro	Tyr	Lys
				275					280					285
Lys	Gln	Lys	Glu	Glu	Gly	Ile	Glu	Asn	Leu	Thr	Asn	Leu	Leu	Asn
				290					295					300
Asp	Asp	Asn	Leu	Ile	Glu	Asn	Asp	Asp	Phe	Ile	Val	Glu	Asp	His
				305					310					315
Pro	Gln	Phe	Asn	Lys	Leu	Ile	Asp	Glu	Ile	Glu	Asn	Ser	Lys	Val
				320					325					330
Gln	Tyr	Leu	Phe	Thr	Met	Gly	Lys	Gly	Gly	Val	Gly	Lys	Thr	Thr
				335					340					345
Val	Ala	Thr	Gln	Leu	Ala	Thr	Ala	Leu	Ser	Asn	Lys	Gly	Tyr	Arg
				350					355					360
Val	Leu	Leu	Ala	Thr	Thr	Asp	Pro	Thr	Lys	Glu	Ile	Asn	Val	Glu
				365					370					375
Thr	Thr	Ser	Asn	Leu	Asn	Thr	Ala	Tyr	Ile	Asp	Glu	Glu	Gln	Ala
				380					385					390
Leu	Glu	Lys	Tyr	Lys	Lys	Glu	Val	Leu	Ala	Thr	Val	Asn	Asp	Asp
				395					400					405
Thr	Pro	Gln	Asp	Asp	Ile	Asp	Tyr	Ile	Met	Glu	Asp	Leu	Lys	Ser
				410					415					420
Pro	Cys	Thr	Glu	Glu	Ile	Ala	Phe	Phe	Lys	Ala	Phe	Ser	Asp	Ile
				425					430					435
Met	Glu	Asn	Gln	Asp	Asp	Met	Asp	Tyr	Val	Ile	Val	Asp	Thr	Ala
				440					445					450
Pro	Thr	Gly	His	Thr	Leu	Leu	Leu	Leu	Asp	Ser	Ser	Glu	Asn	His
				455					460					465
His	Arg	Glu	Leu	Lys	Lys	Lys	Ser	Thr	Gln	Thr	Thr	Ser	Asn	Val

	470		475		480
Glu Thr Leu Leu	Pro Lys Ile Gln Asn	Lys Asn Leu Thr Gln	Met		
	485		490		495
Ile Ile Val Thr	Leu Ala Glu Lys Thr	Pro Tyr Leu Glu Ser	Lys		
	500		505		510
Arg Leu Val Glu	Asp Leu Asn Arg Ala	Asn Ile Gly His Asn	Trp		
	515		520		525
Trp Val Val Asn	Gln Ser Leu Val Thr	Leu Asn Gln Arg Asp	Asp		
	530		535		540
Leu Phe Ser Asn	Lys Lys Glu Asp Glu	Ser Phe Trp Ile Asn	Lys		
	545		550		555
Ile Lys Asn Glu	Ser Phe Asp Asn Tyr	Phe Val Ile Pro Tyr	Gly		
	560		565		570
Gly Leu Ser					

<210> 20  
 <211> 248  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7483850CD1

<400> 20	
Met Asn Ser Asn Leu Asn Leu Asp Gly Phe Leu Leu Pro Ile Ala	
1 5 10 15	
Val Met Asn Ala Ile Ser Ser Leu Pro Leu Leu Ile Leu Ala Pro	
20 25 30	
Phe Leu Glu Tyr Phe Ser Thr Cys Leu Phe Pro Ser Lys Arg Val	
35 40 45	
Gly Ser Phe Leu Ser Thr Cys Ile Ile Ala Gly Asn Leu Phe Ala	
50 55 60	
Ala Leu Ser Val Met Ile Ala Gly Phe Phe Glu Ile His Arg Lys	
65 70 75	
His Phe Pro Ala Val Glu Gln Pro Leu Ser Gly Lys Val Leu Thr	
80 85 90	
Val Ser Ser Met Pro Cys Phe Tyr Leu Ile Leu Gln Tyr Val Leu	
95 100 105	
Leu Gly Val Ala Glu Thr Leu Val Asn Pro Ala Leu Ser Val Ile	
110 115 120	
Ser Tyr Arg Phe Val Pro Ser Asn Val Arg Gly Thr Ser Met Asn	
125 130 135	
Phe Leu Thr Leu Phe Asn Gly Phe Gly Cys Phe Thr Gly Ala Leu	
140 145 150	
Leu Val Lys Leu Val Tyr Leu Ile Ser Glu Gly Lys Asn Arg Gln	
155 160 165	
Trp Phe Pro Asn Thr Leu Asn Lys Gly Asn Leu Glu Gly Phe Phe	
170 175 180	
Phe Phe Leu Ala Ser Leu Thr Leu Leu Asn Val Leu Gly Phe Cys	
185 190 195	
Ser Val Ser Gln Arg Tyr Cys Asn Leu Asn His Phe Asn Ala Gln	
200 205 210	
Asn Ile Arg Gly Ser Asn Leu Glu Glu Thr Leu Leu Leu His Glu	
215 220 225	
Lys Ser Leu Lys Phe Tyr Gly Ser Ile Gln Glu Phe Ser Ser Ser	
230 235 240	
Ile Asp Leu Trp Glu Thr Ala Leu	
245	

<210> 21  
 <211> 761

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5508353CD1

<400> 21

Met	Lys	Pro	Ala	Ser	Pro	Val	Glu	Glu	Glu	Val	Ser	Gln	Val	Cys	
1				5					10					15	
Glu	Ser	Pro	Gln	Cys	Ser	Ser	Ser	Ser	Ala	Cys	Cys	Thr	Glu	Thr	
				20					25					30	
Glu	Lys	Gln	His	Gly	Asp	Ala	Gly	Leu	Leu	Asn	Gly	Lys	Ala	Glu	
				35					40					45	
Ser	Leu	Pro	Gly	Gln	Pro	Leu	Ala	Cys	Asn	Leu	Cys	Tyr	Glu	Ala	
				50					55					60	
Glu	Ser	Pro	Asp	Glu	Ala	Ala	Leu	Val	Tyr	Ala	Ala	Arg	Ala	Tyr	
				65					70					75	
Gln	Cys	Thr	Leu	Arg	Ser	Arg	Thr	Pro	Glu	Gln	Val	Met	Val	Asp	
				80					85					90	
Phe	Ala	Ala	Leu	Gly	Pro	Leu	Thr	Phe	Gln	Leu	Leu	His	Ile	Leu	
				95					100					105	
Pro	Phe	Asp	Ser	Val	Arg	Lys	Arg	Met	Ser	Val	Val	Val	Arg	His	
				110					115					120	
Pro	Leu	Ser	Asn	Gln	Val	Val	Val	Tyr	Thr	Lys	Gly	Ala	Asp	Ser	
				125					130					135	
Val	Ile	Met	Glu	Leu	Leu	Ser	Val	Ala	Ser	Pro	Asp	Gly	Ala	Ser	
				140					145					150	
Leu	Glu	Lys	Gln	Gln	Met	Ile	Val	Arg	Glu	Lys	Thr	Gln	Lys	His	
				155					160					165	
Leu	Asp	Asp	Tyr	Ala	Lys	Gln	Gly	Leu	Arg	Thr	Leu	Cys	Ile	Ala	
				170					175					180	
Lys	Lys	Val	Met	Ser	Asp	Thr	Glu	Tyr	Ala	Glu	Trp	Leu	Arg	Asn	
				185					190					195	
His	Phe	Leu	Ala	Glu	Thr	Ser	Ile	Asp	Asn	Arg	Glu	Glu	Leu	Leu	
				200					205					210	
Leu	Glu	Ser	Ala	Met	Arg	Leu	Glu	Asn	Lys	Leu	Thr	Leu	Leu	Gly	
				215					220					225	
Ala	Thr	Gly	Ile	Glu	Asp	Arg	Leu	Gln	Glu	Gly	Val	Pro	Glu	Ser	
				230					235					240	
Ile	Glu	Ala	Leu	His	Lys	Ala	Gly	Ile	Lys	Ile	Trp	Met	Leu	Thr	
				245					250					255	
Gly	Asp	Lys	Gln	Glu	Thr	Ala	Val	Asn	Ile	Ala	Tyr	Ala	Cys	Lys	
				260					265					270	
Leu	Leu	Glu	Pro	Asp	Asp	Lys	Leu	Phe	Ile	Leu	Asn	Thr	Gln	Ser	
				275					280					285	
Lys	Asp	Ala	Cys	Gly	Met	Leu	Met	Ser	Thr	Ile	Leu	Lys	Glu	Leu	
				290					295					300	
Gln	Lys	Lys	Thr	Gln	Ala	Leu	Pro	Glu	Gln	Val	Ser	Leu	Ser	Glu	
				305					310					315	
Asp	Leu	Leu	Gln	Pro	Pro	Val	Pro	Arg	Asp	Ser	Gly	Leu	Arg	Ala	
				320					325					330	
Gly	Leu	Ile	Ile	Thr	Gly	Lys	Thr	Leu	Glu	Phe	Ala	Leu	Gln	Glu	
				335					340					345	
Ser	Leu	Gln	Lys	Gln	Phe	Leu	Glu	Leu	Thr	Ser	Trp	Gly	Gln	Ala	
				350					355					360	
Val	Val	Cys	Cys	Arg	Ala	Thr	Pro	Leu	Gln	Lys	Ser	Glu	Val	Val	
				365					370					375	
Lys	Leu	Val	Arg	Ser	His	Leu	Gln	Val	Met	Thr	Leu	Ala	Ile	Gly	
				380					385					390	
Asp	Gly	Ala	Asn	Asp	Val	Ser	Met	Ile	Gln	Val	Ala	Asp	Ile	Gly	
				395					400					405	
Ile	Gly	Val	Ser	Gly	Gln	Glu	Gly	Met	Gln	Ala	Val	Met	Ala	Ser	



Asp Phe Ala Val	410	415	420
Ser Gln Phe Lys His	425	430	435
Val His Gly His	440	445	450
Tyr Phe Phe Tyr	455	460	465
Tyr Gln Phe Phe	470	475	480
Trp Val Leu Ile	485	490	495
Val Ile Tyr Gly	500	505	510
Met Gln Leu Pro	515	520	525
Tyr Leu Pro His	530	535	540
Gln Ser Leu Val	545	550	555
Ser Asp Thr Asp	560	565	570
Ala Leu Phe Ile	575	580	585
Leu Thr Trp Ile	590	595	600
Tyr Phe Leu Phe	605	610	615
Asn Pro Pro Ser	620	625	630
Asp Pro Val Phe	635	640	645
Leu Leu Pro Arg	650	655	660
Pro Ser Pro Ile	665	670	675
Glu Glu Arg Thr	680	685	690
Met Asn Gln Val	695	700	705
Ser Gly Arg Arg	710	715	720
Lys Ser Ala Ser	725	730	735
Cys Glu Thr Ala	740	745	750
Glu Met Ala Gly	755	760	

<210> 22

<211> 219

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 8543628CD1

<400> 22

Met Ala Ser Ser Gly Leu Glu Leu Leu Trp Val Ser Leu Pro Gln	
1 5 10 15	
Leu Gly Lys Gly Ala Ala Gln Thr Leu Ser Ile Ser Phe Leu Ser	
20 25 30	
Ile Gly Phe Ser Thr Val Gly Gly Val Leu Tyr Gly Val Leu Arg	

				35				40				45
Thr	Leu	Asn	Asn	Lys	Ala	Ile	Asn	Gly	Val	Leu	Arg	Val
				50				55				60
Glu	Leu	Phe	Arg	Ala	Ile	Pro	Val	Leu	Val	Trp	Leu	Tyr
				65				70				75
Phe	Phe	Gly	Val	Pro	Ile	Phe	Phe	Gly	Leu	Ser	Ile	Pro
				80				85				90
Trp	Cys	Ala	Val	Leu	Val	Leu	Ser	Leu	Trp	Gly	Ala	Ser
				95				100				105
Gly	Glu	Val	Val	Arg	Gly	Ala	Leu	His	Ser	Leu	Pro	Arg
				110				115				120
Arg	Glu	Ala	Gly	Leu	Ser	Ile	Gly	Leu	Ser	Asp	Leu	Gln
				125				130				135
Gly	Tyr	Val	Leu	Leu	Pro	Gln	Ala	Leu	Arg	Arg	Met	Thr
				140				145				150
Thr	Ile	Asn	Val	Tyr	Thr	Arg	Ile	Ile	Lys	Thr	Ser	Ser
				155				160				165
Val	Leu	Ile	Gly	Val	Val	Asp	Val	Ile	Lys	Val	Gly	Gln
				170				175				180
Ile	Glu	Arg	Thr	Tyr	Glu	Ser	Val	Leu	Ile	Tyr	Gly	Ala
				185				190				195
Leu	Phe	Phe	Phe	Phe	Ile	Cys	Tyr	Pro	Leu	Ser	Ala	Ala
				200				205				210
Leu	Leu	Glu	Arg	Arg	Trp	Ala	Gln	Ala				
				215								

<210> 23

<211> 463

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7482754CD1

<400> 23

Met	Glu	Gly	Gln	Thr	Pro	Gly	Ser	Arg	Gly	Leu	Pro	Glu	Lys	Pro
1				5					10					15
His	Pro	Ala	Thr	Ala	Ala	Ala	Thr	Leu	Ser	Ser	Met	Gly	Ala	Val
				20					25					30
Phe	Ile	Leu	Met	Lys	Ser	Ala	Leu	Gly	Ala	Gly	Leu	Leu	Asn	Phe
				35					40					45
Pro	Trp	Ala	Phe	Ser	Lys	Ala	Gly	Gly	Val	Val	Pro	Ala	Phe	Leu
				50					55					60
Val	Glu	Leu	Val	Ser	Leu	Val	Phe	Leu	Ile	Ser	Gly	Leu	Val	Ile
				65					70					75
Leu	Gly	Tyr	Ala	Ala	Ala	Val	Ser	Gly	Gln	Ala	Thr	Tyr	Gln	Gly
				80					85					90
Val	Val	Arg	Gly	Leu	Cys	Gly	Pro	Ala	Ile	Gly	Lys	Leu	Cys	Glu
				95					100					105
Ala	Cys	Phe	Leu	Leu	Asn	Leu	Leu	Met	Ile	Ser	Val	Ala	Phe	Leu
				110					115					120
Arg	Val	Ile	Gly	Asp	Gln	Leu	Glu	Lys	Leu	Cys	Asp	Ser	Leu	Leu
				125					130					135
Ser	Gly	Thr	Pro	Pro	Ala	Pro	Gln	Pro	Trp	Tyr	Ala	Asp	Gln	Arg
				140					145					150
Phe	Thr	Leu	Pro	Leu	Leu	Ser	Val	Leu	Val	Ile	Leu	Pro	Leu	Ser
				155					160					165
Ala	Pro	Arg	Glu	Ile	Ala	Phe	Gln	Lys	Tyr	Thr	Ser	Pro	Ser	His
				170					175					180
Gly	His	Cys	Val	Ser	Ile	Leu	Gly	Thr	Leu	Ala	Ala	Cys	Tyr	Leu
				185					190					195
Ala	Leu	Val	Ile	Thr	Val	Gln	Tyr	Tyr	Leu	Trp	Pro	Gln	Gly	Leu

Val Arg Glu Ser	200	Pro Ser Leu Ser	205	Pro Ala Ser Trp Thr Ser	210
	215		220		225
Val Phe Ser Val	230	Phe Pro Thr Ile Cys	235	Phe Gly Phe Gln Cys His	240
Glu Ala Ala Val	245	Ser Ile Tyr Cys Ser	250	Met Arg Lys Arg Ser Leu	255
Ser His Trp Ala	260	Leu Val Ser Val Leu	265	Ser Leu Leu Ala Cys Cys	270
Leu Ile Tyr Ser	275	Leu Thr Gly Val Tyr	280	Gly Phe Leu Thr Phe Gly	285
Thr Glu Val Ser	290	Ala Asp Val Leu Met	295	Ser Tyr Pro Gly Asn Asp	300
Met Val Ile Ile	305	Val Ala Arg Val Leu	310	Phe Ala Val Ser Ile Val	315
Thr Val Tyr Pro	320	Ile Val Leu Phe Leu	325	Gly Arg Ser Val Met Gln	330
Asp Phe Trp Arg	335	Arg Ser Cys Leu Gly	340	Gly Trp Gly Pro Ser Ala	345
Leu Ala Asp Pro	350	Ser Gly Leu Trp Val	355	Arg Met Pro Leu Thr Ile	360
Leu Trp Val Thr	365	Val Thr Leu Ala Met	370	Ala Leu Phe Met Pro Asp	375
Leu Ser Glu Ile	380	Val Ser Ile Ile Gly	385	Gly Ile Ser Ser Phe Phe	390
Ile Phe Ile Phe	395	Pro Gly Pro Gly Gly	400	Lys His Arg Phe Cys Glu	405
Gly Trp Phe Leu	410	Arg Lys Pro Ser Gln	415	Thr Ala Gln Ala Leu Ser	420
Glu Lys Gly Lys	425	Pro Ala Arg Phe Val	430	Pro His Leu Cys Asn Gly	435
Cys Arg Ala Tyr	440	Arg Thr Lys Ser Gln	445	Val Arg Asn Leu Arg Ala	450
Thr Gly Arg Glu	455	Ser Leu Ser Thr Gln	460	Tyr Glu Gly Ile	

<210> 24  
 <211> 1043  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3794818CD1

<400> 24

Met Glu Phe Val Arg Ala Leu Trp Leu Gly Leu Ala Leu Ala Leu	1	5	10	15
Gly Pro Gly Ser Ala Gly Gly His Pro Gln Pro Cys Gly Val Leu	20	25	30	
Ala Arg Leu Gly Gly Ser Val Arg Leu Gly Ala Leu Leu Pro Arg	35	40	45	
Ala Pro Leu Ala Arg Ala Arg Ala Arg Ala Ala Leu Ala Arg Ala	50	55	60	
Ala Leu Ala Pro Arg Leu Pro His Asn Leu Ser Leu Glu Leu Val	65	70	75	
Val Ala Ala Pro Pro Ala Arg Asp Pro Ala Ser Leu Thr Arg Gly	80	85	90	
Leu Cys Gln Ala Leu Val Pro Pro Gly Val Ala Ala Leu Leu Ala	95	100	105	
Phe Pro Glu Ala Arg Pro Glu Leu Leu Gln Leu His Phe Leu Ala	110	115	120	
Ala Ala Thr Glu Thr Pro Val Leu Ser Leu Leu Arg Arg Glu Ala				

Arg Ala Pro Leu	125	Arg Ala Pro Asn Pro	130	Arg Ala Pro Leu Gln Leu	135
	140		145		150
Trp Ala Ser Pro	155	Leu Glu Thr Leu Leu	160	Asp Val Leu Val Ala	165
	170		175		180
Leu Gln Ala His	185	Ala Trp Glu Asp Val	190	Gly Leu Ala Leu Cys	195
	200		205		210
Thr Gln Asp Pro	215	Gly Gly Leu Val Ala	220	Leu Trp Thr Ser Arg	225
	230		235		240
Gly Arg Pro Pro	245	Gln Leu Val Leu Asp	250	Leu Ser Arg Arg Asp	255
	260		265		270
Gly Asp Ala Gly	275	Leu Arg Ala Arg Leu	280	Ala Pro Met Ala Ala	285
	290		295		300
Val Gly Gly Glu	305	Ala Pro Val Pro Ala	310	Ala Val Leu Leu Gly	315
	320		325		330
Asp Ile Ala Arg	335	Ala Arg Arg Val Leu	340	Glu Ala Val Pro Pro	345
	350		355		360
Pro His Trp Leu	365	Leu Gly Thr Pro Leu	370	Pro Pro Lys Ala Leu	375
	380		385		390
Thr Ala Gly Leu	395	Pro Pro Gly Leu Leu	400	Ala Leu Gly Glu Val	405
	410		415		420
Arg Pro Pro Leu	425	Glu Ala Ala Ile His	430	Asp Ile Val Gln Leu	435
	440		445		450
Ala Arg Ala Leu	455	Gly Ser Ala Ala Gln	460	Val Gln Pro Lys Arg	465
	470		475		480
Leu Leu Pro Ala	485	Pro Val Asn Cys Gly	490	Asp Leu Gln Pro Ala	495
	500		505		510
Pro Glu Ser Pro	515	Gly Arg Phe Leu Ala	520	Arg Phe Leu Ala Asn	525
	530		535		540
Ser Phe Gln Gly	545	Arg Thr Gly Pro Val	550	Val Thr Gly Ser Ser	555
	560		565		570
Gln Val His Met	575	Ser Arg His Phe Lys	580	Val Trp Ser Leu Arg	585
	590		595		600
Asp Pro Arg Gly		Ala Pro Ala Trp Ala		Thr Val Gly Ser Trp	
Asp Gly Gln Leu		Asp Leu Glu Pro Gly		Gly Ala Ser Ala Arg	
Pro Pro Pro Gln		Gly Ala Gln Val Trp		Pro Lys Leu Arg Val	
Thr Leu Leu Glu		His Pro Phe Val Phe		Ala Arg Asp Pro Asp	
Asp Gly Gln Cys		Pro Ala Gly Gln Leu		Cys Leu Asp Pro Gly	
Asn Asp Ser Ala		Thr Leu Asp Ala Leu		Phe Ala Ala Leu Ala	
Gly Ser Ala Pro		Arg Ala Leu Arg Lys		Cys Cys Tyr Gly Tyr	
Ile Asp Leu Leu		Glu Arg Leu Ala Glu		Asp Thr Pro Phe Asp	
Glu Leu Tyr Leu		Val Gly Asp Gly Lys		Tyr Gly Ala Leu Arg	
Gly Arg Trp Thr		Gly Leu Val Gly Asp		Leu Leu Ala Gly Arg	
His Met Ala Val		Thr Ser Phe Ser Ile		Asn Ser Ala Arg Ser	
Val Val Asp Phe		Thr Ser Pro Phe Phe		Ser Thr Ser Leu Gly	
Met Val Arg Ala		Arg Asp Thr Ala Ser		Pro Ile Gly Ala Phe	
Trp Pro Leu His		Trp Ser Thr Trp Leu		Gly Val Phe Ala Ala	
His Leu Thr Ala		Leu Phe Leu Thr Val		Tyr Glu Trp Arg Ser	

Tyr	Gly	Leu	Thr	Pro	Arg	Gly	Arg	Asn	Arg	Ser	Thr	Val	Phe	Ser	
				605					610					615	
Tyr	Ser	Ser	Ala	Leu	Asn	Leu	Cys	Tyr	Ala	Ile	Leu	Phe	Arg	Arg	
				620					625					630	
Thr	Val	Ser	Ser	Lys	Thr	Pro	Lys	Cys	Pro	Thr	Gly	Arg	Leu	Leu	
				635					640					645	
Met	Asn	Leu	Trp	Ala	Ile	Phe	Cys	Leu	Leu	Val	Leu	Ser	Ser	Tyr	
				650					655					660	
Thr	Ala	Asn	Leu	Ala	Ala	Val	Met	Val	Gly	Asp	Lys	Thr	Phe	Glu	
				665					670					675	
Glu	Leu	Ser	Gly	Ile	His	Asp	Pro	Lys	Leu	His	His	Pro	Ala	Gln	
				680					685					690	
Gly	Phe	Arg	Phe	Gly	Thr	Val	Trp	Glu	Ser	Ser	Ala	Glu	Ala	Tyr	
				695					700					705	
Ile	Lys	Lys	Ser	Phe	Pro	Asp	Met	His	Ala	His	Met	Arg	Arg	His	
				710					715					720	
Ser	Ala	Pro	Thr	Thr	Pro	Arg	Gly	Val	Ala	Met	Leu	Thr	Ser	Asp	
				725					730					735	
Pro	Pro	Lys	Leu	Asn	Ala	Phe	Ile	Met	Asp	Lys	Ser	Leu	Leu	Asp	
				740					745					750	
Tyr	Glu	Val	Ser	Ile	Asp	Ala	Asp	Cys	Lys	Leu	Leu	Thr	Val	Gly	
				755					760					765	
Lys	Pro	Phe	Ala	Ile	Glu	Gly	Tyr	Gly	Ile	Gly	Leu	Pro	Gln	Asn	
				770					775					780	
Ser	Pro	Leu	Thr	Ser	Asn	Leu	Ser	Glu	Phe	Ile	Ser	Arg	Tyr	Lys	
				785					790					795	
Ser	Ser	Gly	Phe	Ile	Asp	Leu	Leu	His	Asp	Lys	Trp	Tyr	Lys	Met	
				800					805					810	
Val	Pro	Cys	Gly	Lys	Arg	Val	Phe	Ala	Val	Thr	Glu	Thr	Leu	Gln	
				815					820					825	
Met	Ser	Ile	Tyr	His	Phe	Ala	Gly	Leu	Phe	Val	Leu	Leu	Cys	Leu	
				830					835					840	
Gly	Leu	Gly	Ser	Ala	Leu	Leu	Ser	Ser	Leu	Gly	Glu	His	Ala	Phe	
				845					850					855	
Phe	Arg	Leu	Ala	Leu	Pro	Arg	Ile	Arg	Lys	Gly	Ser	Arg	Leu	Gln	
				860					865					870	
Tyr	Trp	Leu	His	Thr	Ser	Gln	Lys	Ile	His	Arg	Ala	Leu	Asn	Thr	
				875					880					885	
Glu	Pro	Pro	Glu	Gly	Ser	Lys	Glu	Glu	Thr	Ala	Glu	Ala	Glu	Pro	
				890					895					900	
Ser	Gly	Pro	Glu	Val	Glu	Gln	Gln	Gln	Gln	Gln	Gln	Asp	Gln	Pro	
				905					910					915	
Thr	Ala	Pro	Glu	Gly	Trp	Lys	Arg	Ala	Arg	Arg	Ala	Val	Asp	Lys	
				920					925					930	
Glu	Arg	Arg	Val	Arg	Phe	Leu	Leu	Glu	Pro	Ala	Val	Val	Val	Ala	
				935					940					945	
Pro	Glu	Ala	Asp	Ala	Glu	Ala	Glu	Ala	Ala	Pro	Arg	Glu	Gly	Pro	
				950					955					960	
Val	Trp	Leu	Cys	Ser	Tyr	Gly	Arg	Pro	Pro	Ala	Ala	Arg	Pro	Thr	
				965					970					975	
Gly	Ala	Pro	Gln	Pro	Gly	Glu	Leu	Gln	Glu	Leu	Glu	Arg	Arg	Ile	
				980					985					990	
Glu	Val	Ala	Arg	Glu	Arg	Leu	Arg	Gln	Ala	Leu	Val	Arg	Arg	Gly	
				995					1000					1005	
Gln	Leu	Leu	Ala	Gln	Leu	Gly	Asp	Ser	Ala	Arg	His	Arg	Pro	Arg	
				1010					1015					1020	
Arg	Leu	Leu	Gln	Ala	Arg	Ala	Ala	Pro	Ala	Glu	Ala	Pro	Pro	His	
				1025					1030					1035	
Ser	Gly	Arg	Pro	Gly	Ser	Gln	Glu								
				1040											

<210> 25  
<211> 480

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4717525CD1

<400> 25

Met	Arg	Gly	Ser	Pro	Gly	Asp	Ala	Glu	Arg	Arg	Gln	Arg	Trp	Gly	
1				5					10					15	
Arg	Leu	Phe	Glu	Glu	Leu	Asp	Ser	Asn	Lys	Asp	Gly	Arg	Val	Asp	
				20					25					30	
Val	His	Glu	Leu	Arg	Gln	Gly	Leu	Ala	Arg	Leu	Gly	Gly	Gly	Asn	
				35					40					45	
Pro	Asp	Pro	Gly	Ala	Gln	Gln	Gly	Ile	Ser	Ser	Glu	Gly	Asp	Ala	
				50					55					60	
Asp	Pro	Asp	Gly	Gly	Leu	Asp	Leu	Glu	Glu	Phe	Ser	Arg	Tyr	Leu	
				65					70					75	
Gln	Glu	Arg	Glu	Gln	Arg	Leu	Leu	Leu	Met	Phe	His	Ser	Leu	Asp	
				80					85					90	
Arg	Asn	Gln	Asp	Gly	His	Ile	Asp	Val	Ser	Glu	Ile	Gln	Gln	Ser	
				95					100					105	
Phe	Arg	Ala	Leu	Gly	Ile	Ser	Ile	Ser	Leu	Glu	Gln	Ala	Glu	Lys	
				110					115					120	
Ile	Leu	His	Ser	Met	Asp	Arg	Asp	Gly	Thr	Met	Thr	Ile	Asp	Trp	
				125					130					135	
Gln	Glu	Trp	Arg	Asp	His	Phe	Leu	Leu	His	Ser	Leu	Glu	Asn	Val	
				140					145					150	
Glu	Asp	Val	Leu	Tyr	Phe	Trp	Lys	His	Ser	Thr	Val	Leu	Asp	Ile	
				155					160					165	
Gly	Glu	Cys	Leu	Thr	Val	Pro	Asp	Glu	Phe	Ser	Lys	Gln	Glu	Lys	
				170					175					180	
Leu	Thr	Gly	Met	Trp	Trp	Lys	Gln	Leu	Val	Ala	Gly	Ala	Val	Ala	
				185					190					195	
Gly	Ala	Val	Ser	Arg	Thr	Gly	Thr	Ala	Pro	Leu	Asp	Arg	Leu	Lys	
				200					205					210	
Val	Phe	Met	Gln	Val	His	Ala	Ser	Lys	Thr	Asn	Arg	Leu	Asn	Ile	
				215					220					225	
Leu	Gly	Gly	Leu	Arg	Ser	Met	Val	Leu	Glu	Gly	Gly	Ile	Arg	Ser	
				230					235					240	
Leu	Trp	Arg	Gly	Asn	Gly	Ile	Asn	Val	Leu	Lys	Ile	Ala	Pro	Glu	
				245					250					255	
Ser	Ala	Ile	Lys	Phe	Met	Ala	Tyr	Glu	Gln	Ile	Lys	Arg	Ala	Ile	
				260					265					270	
Leu	Gly	Gln	Gln	Glu	Thr	Leu	His	Val	Gln	Glu	Arg	Phe	Val	Ala	
				275					280					285	
Gly	Ser	Leu	Ala	Gly	Ala	Thr	Ala	Gln	Thr	Ile	Ile	Tyr	Pro	Met	
				290					295					300	
Glu	Thr	Leu	Lys	Asn	Trp	Trp	Leu	Gln	Gln	Tyr	Ser	His	Asp	Ser	
				305					310					315	
Ala	Asp	Pro	Gly	Ile	Leu	Val	Leu	Leu	Ala	Cys	Gly	Thr	Ile	Ser	
				320					325					330	
Ser	Thr	Cys	Gly	Gln	Ile	Ala	Ser	Tyr	Pro	Leu	Ala	Leu	Val	Arg	
				335					340					345	
Thr	Arg	Met	Gln	Ala	Gln	Ala	Ser	Ile	Glu	Gly	Gly	Pro	Gln	Leu	
				350					355					360	
Ser	Met	Leu	Gly	Leu	Leu	Arg	His	Ile	Leu	Ser	Gln	Glu	Gly	Met	
				365					370					375	
Arg	Gly	Leu	Tyr	Arg	Gly	Ile	Ala	Pro	Asn	Phe	Met	Lys	Val	Ile	
				380					385					390	
Pro	Ala	Val	Ser	Ile	Ser	Tyr	Val	Val	Tyr	Glu	Asn	Met	Lys	Gln	
				395					400					405	
Ala	Leu	Gly	Val	Thr	Ser	Arg	Leu	Glu	Tyr	Ser	Gly	Ser	Ile	Ser	

Asp His Cys Asn	410	415	420
Leu Cys Leu Pro Gly		Ser Ser Asp Ser Pro	Ala
	425	430	435
Ser Ala Ser Arg Val		Gly Phe His His Val	Ala
	440	445	450
Gln Ala His Leu Gly		Leu Val Gly Ser Arg Asn Ser Ala Ala	Phe
	455	460	465
Ser Leu Pro Thr Cys		Trp Asp Tyr Arg Lys Pro Val Val Met	Pro
	470	475	480

<210> 26  
 <211> 518  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5091793CD1

<400> 26

Met Ala Gly Leu Arg	Asn Glu Ser Glu Gln	Glu Pro Leu Leu Gly
1	5	10
Asp Thr Pro Gly Ser	Arg Glu Trp Asp Ile	Leu Glu Thr Glu Glu
	20	25
His Tyr Lys Ser Arg	Trp Arg Ser Ile Arg	Ile Leu Tyr Leu Thr
	35	40
Met Phe Leu Ser Ser	Val Gly Phe Ser Val	Val Met Met Ser Ile
	50	55
Trp Pro Tyr Leu Gln	Lys Ile Asp Pro Thr	Ala Asp Thr Ser Phe
	65	70
Leu Gly Trp Val Ile	Ala Ser Tyr Ser Leu	Gly Gln Met Val Ala
	80	85
Ser Pro Ile Phe Gly	Leu Trp Ser Asn Tyr	Arg Pro Arg Lys Glu
	95	100
Pro Leu Ile Val Ser	Ile Leu Ile Ser Val	Ala Ala Asn Cys Leu
	110	115
Tyr Ala Tyr Leu His	Ile Pro Ala Ser His	Asn Lys Tyr Tyr Met
	125	130
Leu Val Ala Arg Gly	Leu Leu Gly Ile Gly	Ala Gly Asn Val Ala
	140	145
Val Val Arg Ser Tyr	Thr Ala Gly Ala Thr	Ser Leu Gln Glu Arg
	155	160
Thr Ser Ser Met Ala	Asn Ile Ser Met Cys	Gln Ala Leu Gly Phe
	170	175
Ile Leu Gly Pro Val	Phe Gln Thr Cys Phe	Thr Phe Leu Gly Glu
	185	190
Lys Gly Val Thr Trp	Asp Val Ile Lys Leu	Gln Ile Asn Met Tyr
	200	205
Thr Thr Pro Val Leu	Ser Ala Phe Leu Gly	Ile Leu Asn Ile
	215	220
Ile Leu Ile Leu Ala	Ile Leu Arg Glu His	Arg Val Asp Asp Ser
	230	235
Gly Arg Gln Cys Lys	Ser Ile Asn Phe Glu	Glu Ala Ser Thr Asp
	245	250
Glu Ala Gln Val Pro	Gln Gly Asn Ile Asp	Gln Val Ala Val Val
	260	265
Ala Ile Asn Val Leu	Phe Phe Val Thr Leu	Phe Ile Phe Ala Leu
	275	280
Phe Glu Thr Ile Ile	Thr Pro Leu Thr Met	Asp Met Tyr Ala Trp
	290	295
Thr Gln Glu Gln Ala	Val Leu Tyr Asn Gly	Ile Ile Leu Ala Ala
	305	310

Leu Gly Val Glu Ala Val Val Ile Phe Leu Gly Val Lys Leu Leu	320	325	330
Ser Lys Lys Ile Gly Glu Arg Ala Ile Leu Leu Gly Gly Leu Ile	335	340	345
Val Val Trp Val Gly Phe Phe Ile Leu Leu Pro Trp Gly Asn Gln	350	355	360
Phe Pro Lys Ile Gln Trp Glu Asp Leu His Asn Asn Ser Ile Pro	365	370	375
Asn Thr Thr Phe Gly Glu Ile Ile Ile Gly Leu Trp Lys Ser Pro	380	385	390
Met Glu Asp Asp Asn Glu Arg Pro Thr Gly Cys Ser Ile Glu Gln	395	400	405
Ala Trp Cys Leu Tyr Thr Pro Val Ile His Leu Ala Gln Phe Leu	410	415	420
Thr Ser Ala Val Leu Ile Gly Leu Gly Tyr Pro Val Cys Asn Leu	425	430	435
Met Ser Tyr Thr Leu Tyr Ser Lys Ile Leu Gly Pro Lys Pro Gln	440	445	450
Gly Val Tyr Met Gly Trp Leu Thr Ala Ser Gly Ser Gly Ala Arg	455	460	465
Ile Leu Gly Pro Met Phe Ile Ser Gln Val Tyr Ala His Trp Gly	470	475	480
Pro Arg Trp Ala Phe Ser Leu Val Cys Gly Ile Ile Val Leu Thr	485	490	495
Ile Thr Leu Leu Gly Val Val Tyr Lys Arg Leu Ile Ala Leu Ser	500	505	510
Val Arg Tyr Gly Arg Ile Gln Glu	515		

<210> 27  
 <211> 501  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5945527CD1

<400> 27	
Met Arg Ser Ser Leu Ala Pro Gly Val Trp Phe Phe Arg Ala Phe	1
Ser Arg Asp Ser Trp Phe Arg Gly Leu Ile Leu Leu Leu Thr Phe	20
Leu Ile Tyr Ala Cys Tyr His Met Ser Arg Lys Pro Ile Ser Ile	35
Val Lys Ser Arg Leu His Gln Asn Cys Ser Glu Gln Ile Lys Pro	50
Ile Asn Asp Thr His Ser Leu Asn Asp Thr Met Trp Cys Ser Trp	65
Ala Pro Phe Asp Lys Asp Asn Tyr Lys Glu Leu Leu Gly Gly Val	80
Asp Asn Ala Phe Leu Ile Ala Tyr Ala Ile Gly Met Phe Ile Ser	95
Gly Val Phe Gly Glu Arg Leu Pro Leu Arg Tyr Tyr Leu Ser Ala	110
Gly Met Leu Leu Ser Gly Leu Phe Thr Ser Leu Phe Gly Leu Gly	125
Tyr Phe Trp Asn Ile His Glu Leu Trp Tyr Phe Val Val Ile Gln	140
Val Cys Asn Gly Leu Val Gln Thr Thr Gly Trp Pro Ser Val Val	155
Thr Cys Val Gly Asn Trp Phe Gly Lys Gly Lys Arg Gly Phe Ile	170



Met Gly Ile Trp	Asn Ser His Thr Ser Val Gly Asn Ile Leu Gly	185	190	195
Ser Leu Ile Ala	Gly Ile Trp Val Asn Gly Gln Trp Gly Leu Ser	200	205	210
Phe Ile Val Pro	Gly Ile Ile Thr Ala Val Met Gly Val Ile Thr	215	220	225
Phe Leu Phe Leu	Ile Glu His Pro Glu Asp Val Asp Cys Ala Pro	230	235	240
Pro Gln His His	Gly Glu Pro Ala Glu Asn Gln Asp Asn Pro Glu	245	250	255
Asp Pro Gly Asn	Ser Pro Cys Ser Ile Arg Glu Ser Gly Leu Glu	260	265	270
Thr Val Ala Lys	Cys Ser Lys Gly Pro Cys Glu Glu Pro Ala Ala	275	280	285
Ile Ser Phe Phe	Gly Ala Leu Arg Ile Pro Gly Val Val Glu Phe	290	295	300
Ser Leu Cys Leu	Leu Phe Ala Lys Leu Val Ser Tyr Thr Phe Leu	305	310	315
Tyr Trp Leu Pro	Leu Tyr Ile Ala Asn Val Ala His Phe Ser Ala	320	325	330
Lys Glu Ala Gly	Asp Leu Ser Thr Leu Phe Asp Val Gly Gly Ile	335	340	345
Ile Gly Gly Ile	Val Ala Gly Leu Val Ser Asp Tyr Thr Asn Gly	350	355	360
Arg Ala Thr Thr	Cys Cys Val Met Leu Ile Leu Ala Ala Pro Met	365	370	375
Met Phe Leu Tyr	Asn Tyr Ile Gly Gln Asp Gly Ile Ala Ser Ser	380	385	390
Ile Val Met Leu	Ile Ile Cys Gly Gly Leu Val Asn Gly Pro Tyr	395	400	405
Ala Leu Ile Thr	Thr Ala Val Ser Ala Asp Leu Gly Thr His Lys	410	415	420
Ser Leu Lys Gly	Asn Ala Lys Ala Leu Ser Thr Val Thr Ala Ile	425	430	435
Ile Asp Gly Thr	Gly Ser Ile Gly Ala Ala Leu Gly Pro Leu Leu	440	445	450
Ala Gly Leu Ile	Ser Pro Thr Gly Trp Asn Asn Val Phe Tyr Met	455	460	465
Leu Ile Ser Ala	Asp Val Leu Ala Cys Leu Leu Leu Cys Arg Leu	470	475	480
Val Tyr Lys Glu	Ile Leu Ala Trp Lys Val Ser Leu Ser Arg Gly	485	490	495
Ser Gly Tyr Lys	Glu Ile	500		

<210> 28

<211> 801

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6941124CD1

<400> 28

Met Gln Ala His	Asn Thr Glu Asn Glu Ala Thr Ser Gly Gly Cys	1	5	10	15
Val Leu Leu His	Thr Ser Arg Lys Tyr Leu Lys Leu Lys Asn Phe	20	25	30	
Lys Glu Glu Ile	Arg Ala His Arg Asp Leu Asp Gly Phe Leu Ala	35	40	45	
Gln Ala Ser Ile	Val Leu Asn Glu Thr Ala Thr Ser Leu Asp Asn	50	55	60	

Val	Leu	Arg	Thr	Met	Leu	Arg	Arg	Phe	Ala	Arg	Asp	Pro	Asp	Asn	
				65					70					75	
Asn	Glu	Pro	Asn	Cys	Asn	Leu	Asp	Leu	Leu	Met	Ala	Met	Leu	Phe	
				80					85					90	
Thr	Asp	Ala	Gly	Ala	Pro	Met	Arg	Gly	Lys	Val	His	Leu	Leu	Ser	
				95					100					105	
Asp	Thr	Ile	Gln	Gly	Val	Thr	Ala	Thr	Val	Thr	Gly	Val	Arg	Tyr	
				110					115					120	
Gln	Gln	Ser	Trp	Leu	Cys	Ile	Ile	Cys	Thr	Met	Lys	Ala	Leu	Gln	
				125					130					135	
Lys	Arg	His	Val	Cys	Ile	Ser	Arg	Leu	Val	Arg	Pro	Gln	Asn	Trp	
				140					145					150	
Gly	Glu	Asn	Ser	Cys	Glu	Val	Arg	Phe	Val	Ile	Leu	Val	Leu	Ala	
				155					160					165	
Pro	Pro	Lys	Met	Lys	Ser	Thr	Lys	Thr	Ala	Met	Glu	Val	Ala	Arg	
				170					175					180	
Thr	Phe	Ala	Thr	Met	Phe	Ser	Asp	Ile	Ala	Phe	Arg	Gln	Lys	Leu	
				185					190					195	
Leu	Glu	Thr	Arg	Thr	Glu	Glu	Glu	Phe	Lys	Glu	Ala	Leu	Val	His	
				200					205					210	
Gln	Arg	Gln	Leu	Leu	Thr	Met	Val	Ser	His	Gly	Pro	Val	Ala	Pro	
				215					220					225	
Arg	Thr	Lys	Glu	Arg	Ser	Thr	Val	Ser	Leu	Pro	Ala	His	Arg	His	
				230					235					240	
Pro	Glu	Pro	Pro	Lys	Cys	Lys	Asp	Phe	Val	Pro	Phe	Gly	Lys	Gly	
				245					250					255	
Ile	Arg	Glu	Asp	Ile	Ala	Arg	Arg	Phe	Pro	Leu	Tyr	Pro	Leu	Asp	
				260					265					270	
Phe	Thr	Asp	Gly	Ile	Ile	Gly	Lys	Asn	Lys	Ala	Val	Gly	Lys	Tyr	
				275					280					285	
Ile	Thr	Thr	Thr	Leu	Phe	Leu	Tyr	Phe	Ala	Cys	Leu	Leu	Pro	Thr	
				290					295					300	
Ile	Ala	Phe	Gly	Ser	Leu	Asn	Asp	Glu	Asn	Thr	Asp	Gly	Ala	Ile	
				305					310					315	
Asp	Val	Gln	Lys	Thr	Ile	Ala	Gly	Gln	Ser	Ile	Gly	Gly	Leu	Leu	
				320					325					330	
Tyr	Ala	Leu	Phe	Ser	Gly	Gln	Pro	Leu	Val	Ile	Leu	Leu	Thr	Thr	
				335					340					345	
Ala	Pro	Leu	Ala	Leu	Tyr	Ile	Gln	Val	Ile	Arg	Val	Ile	Cys	Asp	
				350					355					360	
Asp	Tyr	Asp	Leu	Asp	Phe	Asn	Ser	Phe	Tyr	Ala	Trp	Thr	Gly	Leu	
				365					370					375	
Trp	Asn	Ser	Phe	Phe	Leu	Ala	Leu	Tyr	Ala	Phe	Phe	Asn	Leu	Ser	
				380					385					390	
Leu	Val	Met	Ser	Leu	Phe	Lys	Arg	Ser	Thr	Glu	Glu	Ile	Ile	Ala	
				395					400					405	
Leu	Phe	Ile	Ser	Ile	Thr	Phe	Val	Leu	Asp	Ala	Val	Lys	Gly	Thr	
				410					415					420	
Val	Lys	Ile	Phe	Trp	Lys	Tyr	Tyr	Tyr	Gly	His	Tyr	Leu	Asp	Asp	
				425					430					435	
Tyr	His	Thr	Lys	Arg	Thr	Ser	Ser	Leu	Val	Ser	Leu	Ser	Gly	Leu	
				440					445					450	
Gly	Ala	Ser	Leu	Asn	Ala	Ser	Leu	His	Thr	Ala	Leu	Asn	Ala	Ser	
				455					460					465	
Phe	Leu	Ala	Ser	Pro	Thr	Glu	Leu	Pro	Ser	Ala	Thr	His	Ser	Gly	
				470					475					480	
Gln	Ala	Thr	Ala	Val	Leu	Ser	Leu	Leu	Ile	Met	Leu	Gly	Thr	Leu	
				485					490					495	
Trp	Leu	Gly	Tyr	Thr	Leu	Tyr	Gln	Phe	Lys	Lys	Ser	Pro	Tyr	Leu	
				500					505					510	
His	Pro	Cys	Val	Arg	Glu	Ile	Leu	Ser	Asp	Cys	Ala	Leu	Pro	Ile	
				515					520					525	
Ala	Val	Leu	Ala	Phe	Ser	Leu	Ile	Ser	Ser	His	Gly	Phe	Arg	Glu	

Ile Glu Met Ser	530	535	540
Lys Phe Arg Tyr Asn Pro Ser Glu Ser Pro Phe			
545	550	555	
Ala Met Ala Gln Ile Gln Ser Leu Ser Leu Arg Ala Val Ser Gly			
560	565	570	
Ala Met Gly Leu Gly Phe Leu Leu Ser Met Leu Phe Phe Ile Glu			
575	580	585	
Gln Asn Leu Val Ala Ala Leu Val Asn Ala Pro Glu Asn Arg Leu			
590	595	600	
Val Lys Gly Thr Ala Tyr His Trp Asp Leu Leu Leu Leu Ala Ile			
605	610	615	
Ile Asn Thr Gly Leu Ser Leu Phe Gly Leu Pro Trp Ile His Ala			
620	625	630	
Ala Tyr Pro His Ser Pro Leu His Val Arg Ala Leu Ala Leu Val			
635	640	645	
Glu Glu Arg Val Glu Asn Gly His Ile Tyr Asp Thr Ile Val Asn			
650	655	660	
Val Lys Glu Thr Arg Leu Thr Ser Leu Gly Ala Ser Val Leu Val			
665	670	675	
Gly Leu Ser Leu Leu Leu Pro Val Pro Leu Gln Trp Ile Pro			
680	685	690	
Lys Pro Val Leu Tyr Gly Leu Phe Leu Tyr Ile Ala Leu Thr Ser			
695	700	705	
Leu Asp Gly Asn Gln Leu Val Gln Arg Val Ala Leu Leu Leu Lys			
710	715	720	
Glu Gln Thr Ala Tyr Pro Pro Thr His Tyr Ile Arg Arg Val Pro			
725	730	735	
Gln Arg Lys Ile His Tyr Phe Thr Gly Leu Gln Val Leu Gln Leu			
740	745	750	
Leu Leu Leu Cys Ala Phe Gly Met Ser Ser Leu Pro Tyr Met Lys			
755	760	765	
Met Ile Phe Pro Leu Ile Met Ile Ala Met Ile Pro Ile Arg Tyr			
770	775	780	
Ile Leu Leu Pro Arg Ile Ile Glu Ala Lys Tyr Leu Asp Val Met			
785	790	795	
Asp Ala Glu His Arg Pro			
800			

<210> 29

<211> 344

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6972530CD1

<400> 29

Met Ser Ile Phe Leu Leu Leu Arg Met Tyr Ala Ser Ser Leu Phe		
1	5	10
His Arg Phe Gly Trp Ala Gly Leu Ile Val Ala Leu Gly Val His		
20	25	30
Leu Ser Thr Ala Tyr Leu Gly Leu Val Leu Leu Gly Glu Gln His		
35	40	45
Leu Thr Ala Ala Ala Thr Phe Ile Tyr Phe Tyr Leu Thr Thr Thr		
50	55	60
Leu Thr Val Gly Tyr Gly Asp Leu Ala Pro Gln Thr Ser Ala Gly		
65	70	75
Arg Ile Phe Val Ala Ala Trp Val Met Leu Gly Gly Ile Ala Leu		
80	85	90
Leu Thr Ala Ala Ile Gly Lys Thr Thr Ser Ser Val Ile Asp Ala		
95	100	105
Trp Arg Lys Gly Met Lys Gly Lys Gly Asp Phe Thr Gly Lys Val		

	110		115		120
Gly His Thr Val	Leu Ile Gly Trp Glu	Gly Ala Ser Ser Glu	Arg		
	125		130		135
Val Ile Glu Leu	Leu Leu Gln Asp Glu	Thr Ser Asn Asp Asn	Leu		
	140		145		150
Ile Val Ile Cys	Asp Cys Ser Leu Glu	Glu Asn Pro Met Pro	Gly		
	155		160		165
Lys Ala Ala Phe	Ile Arg Gly Glu Ser	Leu Ser Ser Thr Ala	Leu		
	170		175		180
Leu Leu Arg Ala	Gly Val Pro Gly Ala	Glu Arg Val Leu Val	Arg		
	185		190		195
Thr Pro Ser Asp	Asp Leu Thr Leu Ala	Thr Val Leu Ala Val	Asn		
	200		205		210
Gln Leu Ser Pro	Val Gly His Val Val	Ala His Phe Asn Glu	Ser		
	215		220		225
Glu Ile Ala Ala	Leu Ala Ser Ser Tyr	Ala Pro Arg Leu Glu	Cys		
	230		235		240
Thr Ser Ser Met	Ala Ile Glu Met Leu	Val Arg Ala Ser Gln	Asp		
	245		250		255
Pro Gly Ser Ser	Val Val Ile Asn Glu	Leu Leu Cys Val Gly	Gln		
	260		265		270
Gly Ala Thr Gln	Tyr Leu Met Lys Leu	Pro Glu Ala Phe Glu	Ala		
	275		280		285
Thr Phe Gly Glu	Leu Tyr Thr Gln Met	Lys Glu Arg His Asn	Ala		
	290		295		300
Thr Leu Ile Gly	Tyr Arg Ala Lys Gly	Val Gln Gln Pro Ser	Ile		
	305		310		315
Asn Pro Pro Ser	Ala Thr Glu Val Lys	Gly Gly Glu Leu Phe	Tyr		
	320		325		330
Ile Ala Ser Thr	Arg Leu Lys Glu Ile	Ser His Gly Met Ala			
	335		340		

<210> 30

<211> 2701

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6991750CD1

<400> 30

Met His Leu Leu Ile	Ile Tyr Ile Gln Phe	Gln Ile Ala Glu Glu		
1	5	10	15	
Gly Ser Thr Ile Ser	Cys Val Val Glu	Arg Thr Arg Gly Ala	Leu	
	20	25	30	
Asp Tyr Val His Val	Phe Tyr Thr Ile Ser	Gln Ile Glu Thr Asp		
	35	40	45	
Gly Ile Asn Tyr Leu	Val Asp Asp Phe Ala	Asn Ala Ser Gly Thr		
	50	55	60	
Ile Thr Phe Leu Pro	Trp Gln Arg Ser Glu	Val Leu Asn Ile Tyr		
	65	70	75	
Val Leu Asp Asp Asp	Ile Pro Glu Leu Asn	Glu Tyr Phe Arg Val		
	80	85	90	
Thr Leu Val Ser Ala	Ile Pro Gly Asp Gly	Lys Leu Gly Ser Thr		
	95	100	105	
Pro Thr Ser Gly Ala	Ser Ile Asp Pro Glu	Lys Glu Thr Thr Asp		
	110	115	120	
Ile Thr Ile Lys Ala	Ser Asp His Pro Tyr	Gly Leu Leu Gln Phe		
	125	130	135	
Ser Thr Gly Leu Pro	Pro Gln Pro Lys Asp	Ala Met Thr Leu Pro		
	140	145	150	
Ala Ser Ser Val Pro	His Ile Thr Val Glu	Glu Glu Asp Gly Glu		

Ile Arg Leu Leu Val	155	Ile Arg Ala Gln Gly	160	Leu Leu Gly Arg Val	165
Thr Ala Glu Phe Arg	170	Thr Val Ser Leu Thr	175	Ala Phe Ser Pro Glu	180
Asp Tyr Gln Asn Val	185	Ala Gly Thr Leu Glu	190	Phe Gln Pro Gly Glu	195
Arg Tyr Lys Tyr Ile	200	Phe Ile Asn Ile Thr	205	Asp Asn Ser Ile Pro	210
Glu Leu Glu Lys Ser	215	Phe Lys Val Glu Leu	220	Leu Asn Leu Glu Gly	225
Gly Ala Ser Leu Gly	230	Val Ala Ser Gln Ile	235	Leu Val Thr Ile Ala	240
Ala Ser Asp His Ala	245	His Gly Val Phe Glu	250	Phe Ser Pro Glu Ser	255
Leu Phe Val Ser Gly	260	Thr Glu Pro Glu Asp	265	Gly Tyr Ser Thr Val	270
Thr Leu Asn Val Ile	275	Arg His His Gly Thr	280	Leu Ser Pro Val Thr	285
Leu His Trp Asn Ile	290	Asp Ser Asp Pro Asp	295	Gly Asp Leu Ala Phe	300
Thr Ser Gly Asn Ile	305	Thr Phe Glu Ile Gly	310	Gln Thr Ser Ala Asn	315
Ile Thr Val Glu Ile	320	Leu Pro Asp Glu Asp	325	Pro Glu Leu Asp Lys	330
Ala Phe Ser Val Ser	335	Val Leu Ser Val Ser	340	Ser Gly Ser Leu Gly	345
Ala His Ile Asn Ala	350	Thr Leu Thr Val Leu	355	Ala Ser Asp Asp Pro	360
Tyr Gly Ile Phe Ile	365	Phe Ser Glu Lys Asn	370	Arg Pro Val Lys Val	375
Glu Glu Ala Thr Gln	380	Asn Ile Thr Leu Ser	385	Ile Ile Arg Leu Lys	390
Gly Leu Met Gly Lys	395	Val Leu Val Ser Tyr	400	Ala Thr Leu Asp Asp	405
Met Glu Lys Pro Pro	410	Tyr Phe Pro Pro Asn	415	Leu Ala Arg Ala Thr	420
Gln Gly Arg Asp Tyr	425	Ile Pro Ala Ser Gly	430	Phe Ala Leu Phe Gly	435
Ala Asn Gln Ser Glu	440	Ala Thr Ile Ala Ile	445	Ser Ile Leu Asp Asp	450
Asp Glu Pro Glu Arg	455	Ser Glu Ser Val Phe	460	Ile Glu Leu Leu Asn	465
Ser Thr Leu Val Ala	470	Lys Val Gln Ser Arg	475	Ser Ile Pro Asn Ser	480
Pro Arg Leu Gly Pro	485	Lys Val Glu Thr Ile	490	Ala Gln Leu Ile Ile	495
Ile Ala Asn Asp Asp	500	Ala Phe Gly Thr Leu	505	Gln Leu Ser Ala Pro	510
Ile Val Arg Val Ala	515	Glu Asn His Val Gly	520	Pro Ile Ile Asn Val	525
Thr Arg Thr Gly Gly	530	Ala Phe Ala Asp Val	535	Ser Val Lys Phe Lys	540
Ala Val Pro Ile Thr	545	Ala Ile Ala Gly Glu	550	Asp Tyr Ser Ile Ala	555
Ser Ser Asp Val Val	560	Leu Leu Glu Gly Glu	565	Thr Ser Lys Ala Val	570
Pro Ile Tyr Val Ile	575	Asn Asp Ile Tyr Pro	580	Glu Leu Glu Glu Ser	585
Phe Leu Val Gln Leu	590	Met Asn Glu Thr Thr	595	Gly Gly Ala Arg Leu	600
Gly Ala Leu Thr Glu	605	Ala Val Ile Ile Ile	610	Glu Ala Ser Asp Asp	615
	620		625		630

Pro Tyr Gly Leu Phe Gly Phe Gln Ile Thr Lys Leu Ile Val Glu	635	640	645
Glu Pro Glu Phe Asn Ser Val Lys Val Asn Leu Pro Ile Ile Arg	650	655	660
Asn Ser Gly Thr Leu Gly Asn Val Thr Val Gln Trp Val Ala Thr	665	670	675
Ile Asn Gly Gln Leu Ala Thr Gly Asp Leu Arg Val Val Ser Gly	680	685	690
Asn Val Thr Phe Ala Pro Gly Glu Thr Ile Gln Thr Leu Leu Leu	695	700	705
Glu Val Leu Ala Asp Asp Val Pro Glu Ile Glu Glu Val Ile Gln	710	715	720
Val Gln Leu Thr Asp Ala Ser Gly Gly Thr Ile Gly Leu Asp	725	730	735
Arg Ile Ala Asn Ile Ile Ile Pro Ala Asn Asp Asp Pro Tyr Gly	740	745	750
Thr Val Ala Phe Ala Gln Met Val Tyr Arg Val Gln Glu Pro Leu	755	760	765
Glu Arg Ser Ser Cys Ala Asn Ile Thr Val Arg Arg Ser Gly Gly	770	775	780
His Phe Gly Arg Leu Leu Leu Phe Tyr Ser Thr Ser Asp Ile Asp	785	790	795
Val Val Ala Leu Ala Met Glu Glu Gly Gln Asp Leu Leu Ser Tyr	800	805	810
Tyr Glu Ser Pro Ile Gln Gly Val Pro Asp Pro Leu Trp Arg Thr	815	820	825
Trp Met Asn Val Ser Ala Val Gly Glu Pro Leu Tyr Thr Cys Ala	830	835	840
Thr Leu Cys Leu Lys Glu Gln Ala Cys Ser Ala Phe Ser Phe Phe	845	850	855
Ser Ala Ser Glu Gly Pro Gln Cys Phe Trp Met Thr Ser Trp Ile	860	865	870
Ser Pro Ala Val Asn Asn Ser Asp Phe Trp Thr Tyr Arg Lys Asn	875	880	885
Met Thr Arg Val Ala Ser Leu Phe Ser Gly Gln Ala Val Ala Gly	890	895	900
Ser Asp Tyr Glu Pro Val Thr Arg Gln Trp Ala Ile Met Gln Glu	905	910	915
Gly Asp Glu Phe Ala Asn Leu Thr Val Ser Ile Leu Pro Asp Asp	920	925	930
Phe Pro Glu Met Asp Glu Ser Phe Leu Ile Ser Leu Leu Glu Val	935	940	945
His Leu Met Asn Ile Ser Ala Ser Leu Lys Asn Gln Pro Thr Ile	950	955	960
Gly Gln Pro Asn Ile Ser Thr Val Val Ile Ala Leu Asn Gly Asp	965	970	975
Ala Phe Gly Val Phe Val Ile Tyr Asn Ile Ser Pro Asn Thr Ser	980	985	990
Glu Asp Gly Leu Phe Val Glu Val Gln Glu Gln Pro Gln Thr Leu	995	1000	1005
Val Glu Leu Met Ile His Arg Thr Gly Gly Ser Leu Gly Gln Val	1010	1015	1020
Ala Val Glu Trp Arg Val Val Gly Gly Thr Ala Thr Glu Gly Leu	1025	1030	1035
Asp Phe Ile Gly Ala Gly Glu Ile Leu Thr Phe Ala Glu Gly Glu	1040	1045	1050
Thr Lys Lys Thr Val Ile Leu Thr Ile Leu Asp Asp Ser Glu Pro	1055	1060	1065
Glu Asp Asp Glu Ser Ile Ile Val Ser Leu Val Tyr Thr Glu Gly	1070	1075	1080
Gly Ser Arg Ile Leu Pro Ser Ser Asp Thr Val Arg Val Asn Ile	1085	1090	1095
Leu Ala Asn Asp Asn Val Ala Gly Ile Val Ser Phe Gln Thr Ala			

1100	1105	1110
Ser Arg Ser Val Ile	Gly His Glu Gly Glu Ile Leu Gln Phe His	
1115	1120	1125
Val Ile Arg Thr Phe	Pro Gly Arg Gly Asn Val Thr Val Asn Trp	
1130	1135	1140
Lys Ile Ile Gly Gln	Asn Gln Glu Leu Asn Phe Ala Asn Phe Ser	
1145	1150	1155
Gly Gln Leu Phe Phe	Pro Glu Gly Ser Leu Asn Thr Thr Leu Phe	
1160	1165	1170
Val His Leu Leu Asp	Asp Asn Ile Pro Glu Glu Lys Glu Val Tyr	
1175	1180	1185
Gln Val Ile Leu Tyr	Asp Val Arg Thr Gln Gly Val Pro Pro Ala	
1190	1195	1200
Gly Ile Ala Leu Leu	Asp Ala Gln Gly Tyr Ala Ala Val Leu Thr	
1205	1210	1215
Val Glu Ala Ser Asp	Glu Pro His Gly Val Leu Asn Phe Ala Leu	
1220	1225	1230
Ser Ser Arg Phe Val	Leu Leu Gln Glu Ala Asn Ile Thr Ile Gln	
1235	1240	1245
Leu Phe Ile Asn Arg	Glu Phe Gly Ser Leu Gly Ala Ile Asn Val	
1250	1255	1260
Thr Tyr Thr Thr Val	Pro Gly Met Leu Ser Leu Lys Asn Gln Thr	
1265	1270	1275
Val Gly Asn Leu Ala	Glu Pro Glu Val Asp Phe Val Pro Ile Ile	
1280	1285	1290
Gly Phe Leu Ile Leu	Glu Glu Gly Glu Thr Ala Ala Ala Ile Asn	
1295	1300	1305
Ile Thr Ile Leu Glu	Asp Asp Val Pro Glu Leu Glu Glu Tyr Phe	
1310	1315	1320
Leu Val Asn Leu Thr	Tyr Val Gly Leu Thr Met Ala Ala Ser Thr	
1325	1330	1335
Ser Phe Pro Pro Arg	Leu Asp Ser Glu Gly Leu Thr Ala Gln Val	
1340	1345	1350
Ile Ile Asp Ala Asn	Asp Gly Ala Arg Gly Val Ile Glu Trp Gln	
1355	1360	1365
Gln Ser Arg Phe Glu	Val Asn Glu Thr His Gly Ser Leu Thr Leu	
1370	1375	1380
Val Ala Gln Arg Ser	Arg Glu Pro Leu Gly His Val Ser Leu Phe	
1385	1390	1395
Val Tyr Ala Gln Asn	Leu Glu Ala Gln Val Gly Leu Asp Tyr Ile	
1400	1405	1410
Phe Thr Pro Met Ile	Leu His Phe Ala Asp Gly Glu Arg Tyr Lys	
1415	1420	1425
Asn Val Asn Ile Met	Ile Leu Asp Asp Ile Pro Glu Gly Asp	
1430	1435	1440
Glu Lys Phe Gln Leu	Ile Leu Thr Asn Pro Ser Pro Gly Leu Glu	
1445	1450	1455
Leu Gly Lys Asn Thr	Ile Ala Leu Ile Ile Val Leu Ala Asn Asp	
1460	1465	1470
Asp Gly Pro Gly Val	Leu Ser Phe Asn Asn Ser Glu His Phe Phe	
1475	1480	1485
Leu Arg Glu Pro Thr	Ala Leu Tyr Val Gln Glu Ser Val Ala Val	
1490	1495	1500
Leu Tyr Ile Val Arg	Glu Pro Ala Gln Gly Leu Phe Gly Thr Val	
1505	1510	1515
Thr Val Gln Phe Ile	Val Thr Glu Val Asn Ser Ser Asn Glu Ser	
1520	1525	1530
Lys Asp Leu Thr Pro	Ser Lys Gly Tyr Ile Val Leu Glu Glu Gly	
1535	1540	1545
Val Arg Phe Lys Ala	Leu Gln Ile Ser Ala Ile Leu Asp Thr Glu	
1550	1555	1560
Pro Glu Met Asp Glu	Tyr Phe Val Cys Thr Leu Phe Asn Pro Thr	
1565	1570	1575

Gly	Gly	Ala	Arg	Leu	Gly	Val	His	Val	Gln	Thr	Leu	Ile	Thr	Val
				1580					1585					1590
Leu	Gln	Asn	Gln	Ala	Pro	Leu	Gly	Leu	Phe	Ser	Ile	Ser	Ala	Val
				1595					1600					1605
Glu	Asn	Arg	Ala	Thr	Ser	Ile	Asp	Ile	Glu	Glu	Ala	Asn	Arg	Thr
				1610					1615					1620
Val	Tyr	Leu	Asn	Val	Ser	Arg	Thr	Asn	Gly	Ile	Asp	Leu	Ala	Asp
				1625					1630					1635
Leu	Asn	Ile	Glu	Asn	Pro	Lys	Thr	Cys	Glu	Ala	Phe	Asn	Ile	Gly
				1640					1645					1650
Phe	Ser	Pro	Tyr	Phe	Val	Ile	Thr	His	Glu	Glu	Arg	Asn	Glu	Glu
				1655					1660					1665
Lys	Pro	Ser	Leu	Asn	Ser	Val	Phe	Thr	Phe	Thr	Ser	Gly	Phe	Lys
				1670					1675					1680
Leu	Phe	Leu	Val	Gln	Thr	Ile	Ile	Ile	Leu	Glu	Ser	Ser	Gln	Val
				1685					1690					1695
Arg	Tyr	Phe	Thr	Ser	Asp	Ser	Gln	Asp	Tyr	Leu	Ile	Ile	Ala	Ser
				1700					1705					1710
Gln	Arg	Asp	Asp	Ser	Glu	Leu	Thr	Gln	Val	Phe	Arg	Trp	Asn	Gly
				1715					1720					1725
Gly	Ser	Phe	Val	Leu	His	Gln	Lys	Leu	Pro	Val	Arg	Gly	Val	Leu
				1730					1735					1740
Thr	Val	Ala	Leu	Phe	Asn	Lys	Gly	Gly	Ser	Val	Phe	Leu	Ala	Ile
				1745					1750					1755
Ser	Gln	Ala	Asn	Ala	Arg	Leu	Asn	Ser	Leu	Leu	Phe	Arg	Trp	Ser
				1760					1765					1770
Gly	Ser	Gly	Phe	Ile	Asn	Phe	Gln	Glu	Val	Pro	Val	Ser	Gly	Thr
				1775					1780					1785
Thr	Glu	Val	Glu	Ala	Leu	Ser	Ser	Ala	Asn	Asp	Ile	Tyr	Leu	Ile
				1790					1795					1800
Phe	Ala	Lys	Asn	Val	Phe	Leu	Gly	Asp	Gln	Asn	Ser	Ile	Asp	Ile
				1805					1810					1815
Phe	Ile	Trp	Glu	Met	Gly	Gln	Ser	Ser	Phe	Arg	Tyr	Phe	Gln	Ser
				1820					1825					1830
Val	Asp	Phe	Ala	Ala	Val	Asn	Arg	Ile	His	Ser	Phe	Thr	Pro	Ala
				1835					1840					1845
Ser	Gly	Ile	Ala	His	Ile	Leu	Leu	Ile	Gly	Gln	Asp	Met	Ser	Ala
				1850					1855					1860
Leu	Tyr	Cys	Trp	Asn	Ser	Glu	Arg	Asn	Gln	Phe	Ser	Phe	Val	Leu
				1865					1870					1875
Glu	Val	Pro	Ser	Ala	Tyr	Asp	Val	Ala	Ser	Val	Thr	Val	Lys	Ser
				1880					1885					1890
Leu	Asn	Ser	Ser	Lys	Asn	Leu	Ile	Ala	Leu	Val	Gly	Ala	His	Ser
				1895					1900					1905
His	Ile	Tyr	Glu	Leu	Ala	Tyr	Ile	Ser	Ser	His	Ser	Asp	Phe	Ile
				1910					1915					1920
Pro	Ser	Ser	Gly	Glu	Leu	Ile	Phe	Glu	Pro	Gly	Glu	Arg	Glu	Ala
				1925					1930					1935
Thr	Ile	Ala	Val	Asn	Ile	Leu	Asp	Asp	Thr	Val	Pro	Glu	Lys	Glu
				1940					1945					1950
Glu	Ser	Phe	Lys	Val	Gln	Leu	Lys	Asn	Pro	Lys	Gly	Gly	Ala	Glu
				1955					1960					1965
Ile	Gly	Ile	Asn	Asp	Ser	Val	Thr	Ile	Thr	Ile	Leu	Ser	Asn	Asp
				1970					1975					1980
Asp	Ala	Tyr	Gly	Ile	Val	Ala	Phe	Ala	Gln	Asn	Ser	Leu	Tyr	Lys
				1985					1990					1995
Gln	Val	Glu	Glu	Met	Glu	Gln	Asp	Ser	Leu	Val	Thr	Leu	Asn	Val
				2000					2005					2010
Glu	Arg	Leu	Lys	Gly	Thr	Tyr	Gly	Arg	Ile	Thr	Ile	Ala	Trp	Glu
				2015					2020					2025
Ala	Asp	Gly	Ser	Ile	Ser	Asp	Ile	Phe	Pro	Thr	Ser	Gly	Val	Val
				2030					2035					2040
Glu	Lys	Arg	Met	Ser	Ala	Lys	Ile	Leu	Phe	Thr	Glu	Gly	Gln	Val



2045	2050	2055
Leu Ser Thr Ile Thr	Leu Thr Ile Leu Ala Asp Asn Ile Pro Glu	
2060	2065	2070
Leu Ser Glu Val Val	Ile Val Thr Leu Thr Arg Ile Thr Thr Glu	
2075	2080	2085
Gly Val Glu Asp Ser Tyr Lys Gly Ala Thr	Ile Asp Gln Asp Arg	
2090	2095	2100
Ser Lys Ser Val Ile Thr Thr Leu Pro Asn Asp Ser Pro Phe Gly		
2105	2110	2115
Leu Val Gly Trp Arg Ala Ala Ser Val Phe Ile Arg Val Ala Glu		
2120	2125	2130
Pro Lys Glu Asn Thr Thr Thr Leu Gln Leu Gln Ile Ala Arg Asp		
2135	2140	2145
Lys Gly Leu Leu Gly Asp Ile Ala Ile His Leu Arg Ala Gln Pro		
2150	2155	2160
Asn Phe Leu Leu His Val Asp Asn Gln Ala Thr Glu Asn Glu Asp		
2165	2170	2175
Tyr Val Leu Gln Glu Thr Ile Ile Ile Met Lys Glu Asn Ile Lys		
2180	2185	2190
Glu Ala His Ala Glu Val Ser Ile Leu Pro Asp Asp Leu Pro Glu		
2195	2200	2205
Leu Glu Glu Gly Phe Ile Val Thr Ile Thr Glu Val Asn Leu Val		
2210	2215	2220
Asn Ser Asp Phe Ser Thr Gly Gln Pro Ser Val Arg Arg Pro Gly		
2225	2230	2235
Met Glu Ile Ala Glu Ile Met Ile Glu Glu Asn Asp Asp Pro Arg		
2240	2245	2250
Gly Ile Phe Met Phe His Val Thr Arg Gly Ala Gly Glu Val Ile		
2255	2260	2265
Thr Ala Tyr Glu Val Pro Pro Pro Leu Asn Val Leu Gln Val Pro		
2270	2275	2280
Val Val Arg Leu Ala Gly Ser Phe Gly Ala Val Asn Val Tyr Trp		
2285	2290	2295
Lys Ala Ser Pro Asp Ser Ala Gly Leu Glu Asp Phe Lys Pro Ser		
2300	2305	2310
His Gly Ile Leu Glu Phe Ala Asp Lys Gln Val Thr Ala Met Ile		
2315	2320	2325
Glu Ile Thr Ile Ile Asp Asp Ala Glu Phe Glu Leu Thr Glu Thr		
2330	2335	2340
Phe Asn Ile Ser Leu Ile Ser Val Ala Gly Gly Gly Arg Leu Gly		
2345	2350	2355
Asp Asp Val Val Val Thr Val Val Ile Pro Gln Asn Asp Ser Pro		
2360	2365	2370
Phe Gly Val Phe Gly Phe Glu Glu Lys Thr Val Met Ile Asp Glu		
2375	2380	2385
Ser Leu Ser Ser Asp Asp Pro Asp Ser Tyr Val Thr Leu Thr Val		
2390	2395	2400
Val Arg Ser Pro Gly Gly Lys Gly Thr Val Arg Leu Glu Trp Thr		
2405	2410	2415
Ile Asp Glu Lys Ala Lys His Asn Leu Ser Pro Leu Asn Gly Thr		
2420	2425	2430
Leu His Phe Asp Glu Thr Glu Ser Gln Lys Thr Ile Val Leu His		
2435	2440	2445
Thr Leu Gln Asp Thr Val Leu Glu Glu Asp Arg Arg Phe Thr Ile		
2450	2455	2460
Gln Leu Ile Ser Ile Asp Glu Val Glu Ile Ser Pro Val Lys Gly		
2465	2470	2475
Ser Ala Ser Ile Ile Ile Arg Gly Asp Lys Arg Ala Ser Gly Glu		
2480	2485	2490
Val Gly Ile Ala Pro Ser Ser Arg His Ile Leu Ile Gly Glu Pro		
2495	2500	2505
Ser Ala Lys Tyr Asn Gly Thr Ala Ile Ile Ser Leu Val Arg Gly		
2510	2515	2520

Pro Gly Ile Leu Gly Glu Val Thr Val Phe Trp Arg Ile Phe Pro	2525	2530	2535
Pro Ser Val Gly Glu Phe Ala Glu Thr Ser Gly Lys Leu Thr Met	2540	2545	2550
Arg Asp Glu Gln Ser Ala Val Ile Val Val Ile Gln Ala Leu Asn	2555	2560	2565
Asp Asp Ile Pro Glu Glu Lys Ser Phe Tyr Glu Phe Gln Leu Thr	2570	2575	2580
Ala Val Ser Glu Gly Gly Val Leu Ser Glu Ser Ser Ser Thr Ala	2585	2590	2595
Asn Ile Thr Val Val Ala Ser Asp Ser Pro Tyr Gly Arg Phe Ala	2600	2605	2610
Phe Ser His Glu Gln Leu Arg Val Ser Glu Ala Gln Arg Val Asn	2615	2620	2625
Ile Thr Ile Ile Arg Ser Ser Gly Asp Phe Gly His Val Arg Leu	2630	2635	2640
Trp Tyr Lys Thr Met Ser Gly Thr Ala Glu Ala Gly Leu Asp Phe	2645	2650	2655
Val Pro Ala Ala Gly Glu Leu Leu Phe Glu Ala Gly Glu Met Arg	2660	2665	2670
Lys Ser Leu His Val Glu Ile Leu Asp Asp Asp Tyr Pro Glu Gly	2675	2680	2685
Pro Glu Glu Phe Ser Leu Thr Ile Thr Lys Val Glu Leu Gln Gly	2690	2695	2700

Arg

<210> 31  
 <211> 610  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 71726948CD1

<400> 31

Met Asp Thr Pro Arg Gly Ile Gly Thr Phe Val Val Trp Asp Tyr	1	5	10	15
Val Val Phe Ala Gly Met Leu Val Ile Ser Ala Ala Ile Gly Ile	20	25	30	35
Tyr Tyr Ala Phe Ala Gly Gly Gly Gln Gln Thr Ser Lys Asp Phe	40	45	50	55
Leu Met Gly Gly Arg Arg Met Thr Ala Val Pro Val Ala Leu Ser	60	65	70	75
Leu Thr Ala Ser Phe Met Ser Ala Val Thr Val Leu Gly Thr Pro	80	85	90	95
Ser Glu Val Tyr Arg Phe Gly Ala Ile Phe Ser Ile Phe Ala Phe	100	105	110	115
Thr Tyr Phe Phe Val Val Ile Ser Ala Glu Val Phe Leu Pro	120	125	130	135
Val Phe Tyr Lys Leu Gly Ile Thr Ser Thr Tyr Glu Tyr Leu Glu	140	145	150	155
Leu Arg Phe Asn Lys Cys Val Arg Leu Cys Gly Thr Val Leu Phe	160	165	170	175
Ile Val Gln Thr Ile Leu Tyr Thr Gly Ile Val Ile Tyr Ala Pro	180	185	190	195
Ala Leu Ala Leu Asn Gln Val Thr Gly Phe Asp Leu Trp Gly Ala	195			

Met Val Ala Gly Phe Ala Ser Val Ile	Ile Gln Ala Val Val Met	200	205	210
Gln Gly Gly Ile Ser Thr Ile Leu Asn	Asp Ala Tyr Asp Gly Gly	215	220	225
Arg Leu Asn Phe Trp Asn Phe Asn Pro	Asn Pro Leu Gln Arg His	230	235	240
Thr Phe Trp Thr Ile Ile Ile Gly Gly	Thr Phe Thr Trp Thr Ser	245	250	255
Ile Tyr Gly Val Asn Gln Ser Gln Val	Gln Arg Tyr Ile Ser Cys	260	265	270
Lys Ser Arg Phe Gln Ala Lys Leu Ser	Leu Tyr Ile Asn Leu Val	275	280	285
Gly Leu Trp Ala Ile Leu Thr Cys Ser	Val Phe Cys Gly Leu Ala	290	295	300
Leu Tyr Ser Arg Tyr His Asp Cys Asp	Pro Trp Thr Ala Lys Lys	305	310	315
Val Ser Ala Pro Asp Gln Leu Met Pro	Tyr Leu Val Leu Asp Ile	320	325	330
Leu Gln Asp Tyr Pro Gly Leu Pro Gly	Leu Phe Val Ala Cys Ala	335	340	345
Tyr Ser Gly Thr Leu Ser Thr Val Ser	Ser Ser Ile Asn Ala Leu	350	355	360
Ala Ala Val Thr Val Glu Asp Leu Ile	Lys Pro Tyr Phe Arg Ser	365	370	375
Leu Ser Glu Arg Ser Leu Ser Trp Ile	Ser Gln Gly Met Ser Val	380	385	390
Val Tyr Gly Ala Leu Cys Ile Gly Met	Ala Ala Leu Ala Ser Leu	395	400	405
Met Gly Ala Leu Leu Gln Ala Ala Leu	Ser Val Phe Gly Met Val	410	415	420
Gly Gly Pro Leu Met Gly Leu Phe Ala	Leu Gly Ile Leu Val Pro	425	430	435
Phe Ala Asn Ser Ile Gly Ala Leu Val	Gly Leu Met Ala Gly Phe	440	445	450
Ala Ile Ser Leu Trp Val Gly Ile Gly	Ala Gln Ile Tyr Pro Pro	455	460	465
Leu Pro Glu Arg Thr Leu Pro Leu His	Leu Asp Ile Gln Gly Cys	470	475	480
Asn Ser Thr Tyr Asn Glu Thr Asn Leu	Met Thr Thr Thr Glu Met	485	490	495
Pro Phe Thr Thr Ser Val Phe Gln Ile	Tyr Asn Val Gln Arg Thr	500	505	510
Pro Leu Met Asp Asn Trp Tyr Ser Leu	Ser Tyr Leu Tyr Phe Ser	515	520	525
Thr Val Gly Thr Leu Val Thr Leu Leu	Val Gly Ile Leu Val Ser	530	535	540
Leu Ser Thr Gly Gly Arg Lys Gln Asn	Leu Asp Pro Arg Tyr Ile	545	550	555
Leu Thr Lys Glu Asp Phe Leu Ser Asn	Phe Asp Ile Phe Lys Lys	560	565	570
Lys Lys His Val Leu Ser Tyr Lys Ser	His Pro Val Glu Asp Gly	575	580	585
Gly Thr Asp Asn Pro Ala Phe Asn His	Ile Glu Leu Asn Ser Asp	590	595	600
Gln Ser Gly Lys Ser Asn Gly Thr Arg	Leu	605	610	

<210> 32  
 <211> 552  
 <212> PRT  
 <213> Homo sapiens  
 <220>

<221> misc\_feature

<223> Incyte ID No: 7487393CD1

<400> 32

Met Gly Phe Asp Val	Leu Leu Asp Gln Val	Gly Gly Met Gly Arg
1 5	10	15
Phe Gln Ile Cys Leu	Ile Ala Phe Phe Cys	Ile Thr Asn Ile Leu
20 25	30	
Leu Phe Pro Asn Ile	Val Leu Glu Asn Phe	Thr Ala Phe Thr Pro
35 40	45	
Ser His Arg Cys Trp	Val Pro Leu Leu Asp	Asn Asp Thr Val Ser
50 55	60	
Asp Asn Asp Thr Gly	Thr Leu Ser Lys Asp	Asp Leu Leu Arg Ile
65 70	75	
Ser Ile Pro Leu Asp	Ser Asn Leu Arg Pro	Gln Lys Cys Gln Arg
80 85	90	
Phe Ile His Pro Gln	Trp Gln Leu Leu His	Leu Asn Gly Thr Phe
95 100	105	
Pro Asn Thr Asn Glu	Pro Asp Thr Glu Pro	Cys Val Asp Gly Trp
110 115	120	
Val Tyr Asp Arg Ser	Ser Phe Leu Ser Thr	Ile Val Thr Glu Trp
125 130	135	
Asp Leu Val Cys Glu	Ser Gln Ser Leu Lys	Ser Met Val Gln Ser
140 145	150	
Leu Phe Met Ala Gly	Ser Leu Leu Gly Gly	Leu Ile Tyr Gly His
155 160	165	
Leu Ser Asp Arg Val	Gly Arg Lys Ile Ile	Cys Lys Leu Cys Phe
170 175	180	
Leu Gln Leu Ala Ile	Ser Asn Thr Cys Ala	Ala Phe Ala Pro Thr
185 190	195	
Phe Leu Val Tyr Cys	Ile Leu Arg Phe Leu	Ala Gly Phe Ser Thr
200 205	210	
Met Thr Ile Leu Gly	Asn Thr Phe Ile Leu	Ser Leu Glu Trp Thr
215 220	225	
Leu Pro Arg Ser Arg	Ser Met Thr Ile Met	Val Leu Leu Cys Ser
230 235	240	
Tyr Ser Val Gly Gln	Met Leu Leu Gly Gly	Leu Ala Phe Ala Ile
245 250	255	
Gln Asp Trp His Ile	Leu Gln Leu Thr Val	Ser Thr Pro Ile Ile
260 265	270	
Val Leu Phe Leu Ser	Ser Trp Lys Met Val	Glu Ser Ala Arg Trp
275 280	285	
Leu Ile Ile Asn Asn	Gln Leu Asp Glu Gly	Leu Lys Glu Leu Arg
290 295	300	
Arg Val Ala His Ile	Asn Gly Lys Lys Asn	Thr Glu Glu Thr Leu
305 310	315	
Thr Thr Glu Leu Val	Arg Ser Thr Met Lys	Lys Glu Leu Asp Ala
320 325	330	
Val Arg Ile Lys Thr	Ser Ile Phe Ser Leu	Phe Arg Ala Pro Lys
335 340	345	
Leu Arg Met Arg Val	Phe Gly Leu Cys Phe	Val Arg Phe Ala Ile
350 355	360	
Thr Val Pro Phe Tyr	Gly Leu Ile Leu Asn	Leu Gln His Leu Gly
365 370	375	
Ser Asn Val Ser Leu	Phe Gln Ile Leu Cys	Gly Ala Val Thr Phe
380 385	390	
Thr Ala Arg Cys Val	Ser Leu Leu Thr Leu	Asn His Met Gly Arg
395 400	405	
Arg Ile Ser Gln Ile	Leu Phe Thr Phe Pro	Val Gly Leu Phe Ile
410 415	420	
Leu Val Asn Thr Phe	Leu Pro Gln Glu Met	Gln Ile Leu Arg Val
425 430	435	
Val Leu Ala Thr Leu	Gly Ile Gly Ser Val	Ser Ala Ala Ser Asn

	440		445		450
Ser Ala Ser Val	His His Asn Glu Leu	Val Pro Thr Ile Leu	Arg		
	455		460		465
Ser Thr Val Ala	Gly Ile Asn Ala Val	Ser Gly Arg Thr Gly	Ala		
	470		475		480
Ala Leu Ala Pro	Leu Leu Met Thr Leu	Met Ala Tyr Ser Pro	His		
	485		490		495
Leu Pro Trp Ile	Ser Tyr Gly Val Phe	Pro Ile Leu Ala Val	Pro		
	500		505		510
Val Ile Leu Leu	Leu Pro Glu Thr Arg	Asp Leu Pro Leu Pro	Asn		
	515		520		525
Thr Ile Gln Asp	Val Glu Asn Asp Arg	Lys Asp Ser Arg Asn	Ile		
	530		535		540
Lys Gln Glu Asp	Thr Cys Met Lys Val	Thr Gln Phe			
	545		550		

<210> 33

<211> 2365

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7484831CB1

<400> 33

```

atggagcctg gtgacgcggc gcgccctggc tcgggtcggg ctaccggggc gccgcgcgcg 60
cggctgctgc tgctgccgct gctgctgggt tgggggctgc gagtcgcggc cgcggcctcg 120
gcctcctcct ctggggcgcc ggccgaggac agcagcgcca tggaggagct cgctactgag 180
aaggaggcgg aggagagcca ccggcaagac agcgtgagcc tgctcacctt catcctgctg 240
ctcacgctca ccatcctcac catctggctc ttcaagcacc gccgggtgcg ctttctgcac 300
gagaccgggc tggccatgat ctatgggctc atcgttgggg tgatcctgag gtatggtacc 360
cctgctacca gtggccgtga caaatcactc agctgcactc aggaagacag ggccttcagt 420
accttattag tgaatgtcag cggaaagtgc ttccaataca ctctgaaagg agaaatcagt 480
cctggcaaga tcaacagcgt agagcagaat gatatgtac ggaaggtaac attcgatcca 540
gaagtatttt tcaacattct tctgcctcca attatttttc atgctggata cagcttaaag 600
aagagacact ttttcagaaa tcttggatct atactggcct atgccttctt ggggactgct 660
gtttcatgct tcattattgg aaatctcatg tatggtgtgg tgaagctcat gaagattatg 720
ggacagctct cagataaatt ttactacaca gattgtctct tttttggagc aatcatctct 780
gccactgacc cagtgactgt gctggcgata tttaataaat tgcattgcaga cgtggatctt 840
tacgcacttc tttttggaga gagcgctcta aatgatgctg ttgccattgt actgtcctcg 900
tctattgttg cctaccagcc agcgggactg aacactcacg cctttgatgc tgctgccttt 960
tttaagtcag ttggcatttt tctaggtata ttttagtggc cttttaccat gggagctgtg 1020
actggtgttg tgactgctct agtgactaag tttaccaaac tgcactgctt cccctgctg 1080
gagacggcgc tgttcttctc catgtcctgg agcacgtttc tcttggcaga agcctgcgga 1140
tttacaggtg ttgtagctgt ccttttctgt ggaatcacac aagctcata cacctacaac 1200
aatctgtcgg tggaaatcaag aagtccaacc aagcagctct ttgaggtgtt acatttctg 1260
gcagagaact tcattctctc ctacatgggc ctggcactgt ttaccttcca gaagcacgtt 1320
ttcagcccca ttttcatcat cggagctttt gttgccatct tcttgggcag agccgcgcac 1380
atctaccgcg tctccttctt cctcaacttg ggcagaaggc ataagattgg ctggaatttt 1440
caacacatga tgatgttttc aggcctcagg ggagcaatgg catttgcgtt ggccatccgt 1500
gacacggcat cctatgctcg ccagatgatg ttcaagacca ccttctcat tgtgttcttc 1560
actgtctgga tcattggagg aggcacgaca cccatgttgt catggcttaa catcagagt 1620
ggcgtcgagg agccctccga agaggaccag aatgaacacc actggcagta cttcagagt 1680
ggtgttgacc ccgatcaaga cccaccaccc aacaacgaca gctttcaagt cttacaagg 1740
gacggccagc attctgccag aggaaccggg acaaaacagg agagcacatg gatattcagg 1800
ctgtggtaca gctttgatca caattatctg aagcccatcc tcacacacag tgggtcccca 1860
ctaaccacca cgtccccgcg ctggtgtggc ttactagctc gatgtctgac cagtccccag 1920
gtgtacgata accaagagcc actgagagag gaagactctg atttcatcct gaccgaaggc 1980
gacctgacat tgacctacgg ggacagcaca gtgactgcaa atggctctc aagttcgcac 2040
accgcctcca cgagtctgga gggcagccgg agaacgaaga gcagctcgga ggaagtgtg 2100
gagcgagacc tgggaatggg agaccagaag gtttcgagcc ggggcacccg cctagtgttt 2160
ccccgggaag ataatgcttg actttcccc caagccctgg cgcatggga ctgaagattg 2220

```

aaaatacatc	cacgaacatt	tcaacatgga	acgaagaatt	atagttcctt	ctctggctat	2280
atccataaag	aagtagttat	gaaatgttta	aaaccaaagg	caaatagtgt	tataactctta	2340
tttttgatta	atctgaggaa	aggag				2365

<210> 34  
 <211> 3400  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2477266CB1

<400> 34

agcggcgct	cagttgctgc	catggagtca	gccgggctag	agcagctcct	acgggagctg	60
ctgctaccgg	acaccgagcg	catccgtcgg	gccacggaac	agctccagat	cgttcttcgg	120
gccccgcg	ctttgcggc	tctctgcgac	ctgctagcct	cggcggccga	ccccagatc	180
cgccagtttg	cggccgtgct	gaccgcgaga	cgactgaaca	cccgtggcg	acggctggcg	240
gcgagcaac	gggagagcct	caagtccttg	atcctgacgg	ccctgcagag	agaaacagag	300
cactgtgtga	gcctcagcct	ggcccagctc	tcagccacca	ttttcgaaa	ggaaggcttg	360
gaggcctggc	cacagctttt	gcagctgctt	cagcacagta	cccacagccc	ccacagccca	420
gagagagaga	tggggctttt	gctgctaagt	gtggtggtga	cctcccggcc	cgaggccttc	480
caacccacc	accgggagct	tcttcggctt	ctgaatgaga	ctcttggtga	ggtgggctct	540
cctgggctgc	tcttctactc	cctgcgcact	ctgaccacca	tggtcccta	cctcagcact	600
gaagatgtgc	ctctcgctcg	gatgttggtg	ccaagctga	tcattggccat	gcagactctg	660
atccccatag	atgaggcaaa	ggcctgtgag	gcccttgagg	ctttggatga	actgttgagg	720
tcagaggtgc	cggctcatcac	cccctacctc	tctgaagtc	tcacattctg	cctggaggta	780
gctagaaatg	tggccctggg	caatgcgata	cgcatacgta	ttctctgctg	cctcactttc	840
ttggtcaaa	tcaagagcaa	ggccttactg	aagaatcgte	tcttgccacc	cttgcctgcac	900
acccttttcc	ccattgtggc	tgctgagccc	ccaccaggcc	agttggatcc	cgaggaccag	960
gattcagaag	aggaagaggt	ggagattgag	ctgatggggg	agactcccaa	gcatttcgct	1020
gtacaagttg	tggacatgct	ggcactacac	ctgcccccg	agaagctctg	tccccagctg	1080
atgcccatgt	tggaaagaggc	tttgcggagc	gagagcccat	accagcgcaa	agctggactc	1140
ctggtgctgg	ccgtgctgtc	tgacggagct	ggcgaccaca	tcaggcagag	actgctgccc	1200
ccactgttgc	agattgtgtg	caagggcctg	gaggaccctc	cgcaagttgt	acgcaatgct	1260
gcgctgtttg	ccctgggcca	gttctcagaa	aacctacagc	cccatatcag	cagctattca	1320
agggaggtaa	tgccactgct	cctgccttac	ttgaagtcgg	tgctctcttg	acacacacac	1380
cacctagcca	aggcctgcta	tgccctggag	aattttgtgg	agaacctagg	gccccagggtg	1440
cagccctacc	ttccggagct	tatggaatgc	atgctgcagc	ttctgaggaa	ccccagcagt	1500
ccccgggcca	aggagctggc	tgtgagcgcc	ctgggagcca	ttgctacggc	tgcccaggcc	1560
tcgctgtctg	cctacttccc	tgccatcatg	gagcacctgc	gggaattcct	gttaacaggc	1620
cgtgaggacc	ttcagcctgt	gcagatccag	agcctggaga	cactgggggt	gctggcacga	1680
gcagtggggg	agcccatgag	gccgctggct	gaggaatgct	gccagctggg	tctgggcctc	1740
tgcgaccagg	tagacgacgc	tgacttgccg	cgtgcacgt	acagcctatt	tgcagcctta	1800
tcgggtctga	tgggtgaggg	cctggcgccc	cacttggaac	agatcaccac	gctcatgctg	1860
ctgtcactgc	gttccaccga	gggcatttgt	cctcagtatg	acgggagcag	ctccttccct	1920
ctgtttgacg	atgagagtga	tggggaagaa	gaggaggagc	tcattggatga	ggatgtggaa	1980
gaagaggatg	actcagagat	ctcagggtac	agcgtggaga	atgccttctt	cgatgagaag	2040
gaagacacct	gtgctgccgt	gggggagatc	tctgtgaaca	ccagtgtggc	cttccttcca	2100
tacatggaaa	gtgtctttga	agaagtattt	aaactgctgg	agtgcctca	cctgaatgtg	2160
cggaaggcag	cccatgagge	tctgggtcag	ttttgctgtg	cactgcacaa	ggcctgtcaa	2220
agctgcccc	cggaaaccaa	cactgctgct	ttgcaggctg	ccctggcccc	agtcgtgcca	2280
tcctacatgc	aggcagtga	cagggagcgg	gaacgccagg	tggtgatggc	cgtgctggag	2340
gccctgacag	gggtgctccg	cagctgtggg	accctcacac	tgaagcccc	tgggcgcctc	2400
gctgagctct	gtggcgtgct	caaggctgtg	ctgcagagga	agacagcctg	tcaggatact	2460
gacgaggagg	aggaagagga	agatgatgat	caggctgaat	acgacgccat	gttgctggag	2520
cacgctggag	agggcatccc	tgccctggca	gccgcggctg	ggggagactc	ctttgcccc	2580
ttctttgccg	gtttcctgcc	attattgggtg	tgcaagacaa	aacagggtctg	cacagtggca	2640
gagaagtcc	ttgcagtggg	gaccttggca	gagactatc	agggcctggg	tgctgcctca	2700
gccagtttg	tgtctcggt	gctccctgtg	ctgtttagca	ccgcccaaga	ggcagacccc	2760
gaggtgcgaa	gcaatgccat	cttcgggatg	ggcgtgctgg	cagagcatgg	gggccacct	2820
gccagggaac	acttccccaa	gctgctgggg	ctccttttct	cctcctggc	gcgggagcga	2880
catgatcgtg	tccgtgacaa	catctgtggg	gcacttgccc	gcctgttgat	ggccagtcct	2940

accaggaaac	cagagcccca	ggtgctggct	gccctactgc	atgccctgcc	actgaaggag	3000
gacttggagg	agtgggtcac	cattgggcgc	ctcttcagct	tectgtacca	gagcagccct	3060
gaccaggtta	tagatgtggc	tcccagagctt	ctgcgtatct	gcagcctcat	tctggctgac	3120
aacaagatcc	caccagacac	caaggccgca	ctgttgcctg	tcctgacgtt	cctggccaaa	3180
cagcacaccg	acagctttca	agcagctctg	ggctcactgc	ctgttgacaa	ggctcaggag	3240
ctccaggctg	tactgggcct	ctcctagact	gcaggctgca	gccagtccag	agagaataga	3300
gcctgcccag	gccttaagac	cacctctcag	cccagttcag	ttctgcctta	ccaaagattc	3360
tgagactcat	accatttggg	agccagcccc	acttgcctgc			3400

<210> 35

<211> 4458

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3552033CB1

<400> 35

gttgttgcta	gtatgatata	gatcatgtgt	acgtctatgt	tagtgctatg	tacgttgtta	60
tagtgtcttg	cttgttgtat	ctatatgtgt	attgtttcgt	tttgtgtgtg	tatatattat	120
tatcgtttgc	cccgtttaat	tacctcccat	ggggcccccga	gttttatttt	cccccttgg	180
tgggatgcgt	cttcaatcac	cgcttttata	ccctgatatt	cggttcaagg	gactttctcc	240
gggggctctt	tgggcaccac	cgcaagaacc	ggtctcagct	caaatccccg	aatgaaggcc	300
tttaaaaaga	agatgtgccc	aacagtgaag	ccccgcccc	aggcccccata	ggtttccttg	360
ggcttctgat	agggccagca	ttggccctca	ccccagtat	ctccggcttc	agaattgaga	420
aaaatgaaag	tgcctctctc	cgaagggaca	tccagtctga	aaagtgggtcc	atcagcaaac	480
acactcaact	cagccctacg	gatgcttttg	ggaccattga	gttccaaggga	ggtggccatt	540
ccaacaaagc	catgtatgtg	cgagtatctt	ttgatacaaa	acctgatctc	ctettacacc	600
tgatgaccaa	ggaatggcag	ttggagcttc	ccaagcttct	catctctgtc	catgggggcc	660
tgcagaactt	tgaactccag	ctgaaactca	agcaagtott	tgggaaaggg	ctcatcaaa	720
cagctatgac	aactggagcg	tggatattca	ctggaggggt	taacacaggt	gttattcgtc	780
atgtttggcg	tgccttgaag	gatcatgcct	ctaagtctcg	aggaaagata	tgcaccatag	840
gtattgcccc	ctggggaatt	gtggaaaacc	aggaggacct	cattggaaga	gatgttgctc	900
ggccatacca	gaccatgtcc	aatcccatga	gcaagctcac	tgttctcaac	agcatgcatt	960
cccacttcat	tctggctgac	aacgggacca	ctggaaaata	tggagcagag	gtgaaacttc	1020
gaagacaact	ggaaaagcat	atttcaactc	agaagataaa	cacaagaatc	ggtcaagggtg	1080
ttcctgtggg	ggcactcata	gtggaaggag	gacccaatgt	gatctcgatt	gttttggagt	1140
accttcgaga	cacccctccc	gtgccagtgg	ttgtctgtga	tgggagtggg	cgggcatcgg	1200
acatcctggc	ctttgggcat	aaatactcag	aagaaggcgg	actgataaat	gaatctttga	1260
gggaccagct	gttggtagct	atacagaaga	ctttcacata	cactcgaacc	caagctcagc	1320
atctgttcat	catcctcatg	gagtgcatga	agaagaagga	attgattacg	gtatttcgga	1380
tgggatcaga	aggacaccag	gacattgatt	tggctatcct	gacagcttta	ctcaaaggag	1440
ccaatgcctc	ggccccagac	caactgagct	tagctttagc	ctggaacaga	gtcgacatcg	1500
ctcgagcca	gatctttatt	tacgggcaac	agtggccggg	gggatctctg	gagcaagcca	1560
tgttggtatg	cttagttctg	gacagagtgg	attttgtgaa	attactcata	gagaatggag	1620
taagcatgca	ccgttttctc	accatctcca	gactagagga	attgtacaat	acgagacatg	1680
ggccctcaaa	tacattgtac	cacttgggtc	gggatgtcaa	aaaggggaac	ctgccccag	1740
actacagaat	cagcctgatt	gacatcggcc	tgggtgatcg	gtacctgatg	ggcggggctt	1800
atcgctgcaa	ctacacgcgc	aagcgcttcc	ggaccctcta	ccacaacctc	ttcgccccc	1860
agaggcccaa	agccttgaaa	ctgctgggaa	tggaggatga	tattcccttg	aggcgaggaa	1920
gaaagacaac	caagaaacgt	gaagaagagg	tggacattga	cttggatgat	cctgagatca	1980
accacttccc	cttcccttcc	catgagctca	tgggtgtggc	tgttctcatg	aagcggcaga	2040
agatggccct	gttcttctgg	cagcacgggtg	aggaggccat	ggccaaggcc	ctggtggcct	2100
gcaagctctg	caaagccatg	gctcatgagg	cctctgagaa	cgacatgggt	gacgacattt	2160
cccaggagct	gaatcacaat	tccagagact	ttggccagct	ggctgtggag	ctcctggacc	2220
agtcctacaa	caggacgaa	cagctggcca	tgaactgct	gacgtatgag	ctgaagaact	2280
ggagcaacgc	cacgtgcctg	cagcttgccg	tggctgcca	acaccgcgac	ttcatcgcg	2340
acacgtgcag	ccagatgctg	ctcaccgaca	tgtggatggg	ccggctccgc	atgcgcaaga	2400
actcaggcct	caaggttaatt	ctgggaattc	tacttctctc	ttcaattctc	agcttggagt	2460
tcaagaacaa	agacgacatg	ccctatatgt	ctcaggccca	ggaaatccac	ctccaagaga	2520
aggaggcaga	agaaccagag	agcccacaa	aggaaaaaga	ggaagaggac	atggagctca	2580
tagcaatggt	gggacgaaac	aacggggagt	cctccaggaa	gaaggatgaa	gaggaagtcc	2640

agagcgcagca	ccggttaate	cccctcggca	gaaaaatcta	tgaattctac	aatgcaccca	2700
tcgtgaagtt	ctggttctac	acactggcgt	atateggata	cctgatgctc	ttcaactata	2760
tcgtgtttagt	gaagatggaa	cactggccgc	cccccagga	atggatcgta	atctcctata	2820
ttttcacccct	gggaatagaa	aagatgagag	agattctgat	gtcagagcca	gggaagtgtc	2880
tacagaaagt	gaaggtatgg	ctgcaggagc	actggaatgt	cacggacctc	atcgccatcc	2940
ttctgttttc	tgtcgggaatg	atccttcgtc	tccaagacca	gcccttcagg	agtgcaggga	3000
gggtcatcta	ctgcgtgaac	atcatttact	ggtatatccg	tctcctagac	atcttcggcg	3060
tgaacaagta	tttgggccc	tatgtaatga	tgattggaaa	aatgatgata	gacatgatgt	3120
actttgtcat	cattatgctg	gtggttctga	tgagctttgg	ggtcgccagg	caagccatcc	3180
tttttcccaa	tgaggagcca	tcatggaaac	tggccaagaa	catcttctac	atgccctatt	3240
ggatgattta	tggggaagt	tttgcggacc	agatagaccc	tccctgtgga	cagaatgaga	3300
cccgcagagga	tggtaaaata	atccagctgc	ctccctgcaa	gacaggagct	tggatcgtgc	3360
cggccatcat	ggcctgctac	ctcttagtgg	caaacatctt	gctggtcaac	ctcctcattg	3420
ctgtctttta	caatacattt	tttgaagtaa	aatcgatata	caaccaagtc	tgggaagtctc	3480
agaggtatca	gctcatcatg	actttccatg	aaaggccagt	tctgccccca	ccactgatca	3540
tcttcagcca	catgaccatg	atattccagc	acctgtgctg	ccgatggagg	aaacacgaga	3600
gcgacccgga	tgaaggggac	tacggcctga	aactcttcat	aaccgatgat	gagctcaaga	3660
aagtacatga	ctttgaagag	caatgcatag	aagaatactt	cagagaaaag	gatgatcggt	3720
tcaactcatc	taatgatgag	aggatacggg	tgacttcaga	aaggggtggag	aacatgtcta	3780
tgcggtctgga	ggaagtcaac	gagagagagc	actccatgaa	ggcttcaactc	cagaccgtgg	3840
acatccggct	ggcgagctg	gaagacctta	tcgggcgcac	ggccacggcc	ctggagcgcc	3900
tgacaggtct	ggagcgggcc	gagtccaaca	aaatccgctc	gaggacctcg	tcagactgca	3960
gggacgccc	ctacattgtc	cgtcagagca	gcttcaacag	ccaggaagg	aacaccttca	4020
agctccaaga	gagtatagac	cctgcagagc	atcctcttta	ttcagtataa	gccggcagca	4080
agcagttcta	cctaacgtcc	cacatccttc	tcatgccaac	acttctgtaa	ttgatcatta	4140
taaagaaaaa	acaaggtaac	agtcatagtt	cacctgtctc	ttatctattc	acttctgggtg	4200
ccacaactgt	ttatcctttt	ttgaagaaaa	taagggaaaca	gaaatgcctt	tttgtattgc	4260
aatcgaaatg	aaagaagagt	tgatgttaaa	aaaacaaaag	tcaagtgatt	tattatatac	4320
agtgggcggt	caagtctagt	cgagcaagct	caggagaatg	taattaaata	attttatatt	4380
ttttaattta	ttttgtatct	cacctgtcat	ggatgaattc	attcactgaa	tatgtaatat	4440
tgaaccttaa	aaaaaaaa					4458

<210> 36

<211> 2722

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4778139CB1

<220>

<221> unsure

<222> 2609

<223> a, t, c, g, or other

<400> 36

atggagatcg	ctggagagaa	gaaataccag	tattatagca	tgccctcttc	tttgetgagg	60
agcctggctg	gaggccggca	cccccaaga	gtgcagagg	gggcacttct	tgagattaca	120
aacagtaaaa	gagaggctac	aaatgtaaga	aatgaccagg	aaaggcaaga	aacacaaagt	180
agtatagtgg	tttctgggg	gtctcctaac	aggcaagcac	actcaaagta	tggccagttt	240
cttctgggtcc	cctctaattct	aaagcgagtt	cctttttcag	cagaaactgt	cttgccctctg	300
tccagaccct	ctgtgccaga	tgtgctggca	actgaacagg	acatccagac	tgaggttctt	360
gttcatctga	ctgggcagac	cccagttgtc	tctgactggg	catcagtggg	tgaacccaag	420
gaaaagcacg	agcctattgc	tcaattactg	gatggacaag	acaaggcaga	gcaagtgcta	480
cccactttga	gttgacacac	tgaacccatg	acaatgagct	cccctctttc	ccaagccaag	540
atcatgcaaa	ctggaggtgg	atatgtaaac	tgggcatttt	cagaagggtga	tgaaactggg	600
gtgttttagca	tcaagaaaaa	gtggcaaac	tgcttgccct	ccacttggtga	cagtgattcc	660
tctcggagtg	aacagcacca	gaagcaggcc	caggacagct	ccctatctga	taactcaaca	720
agatcggccc	agagtagtga	atgctcagag	gtgggacat	ggcttcagcc	aaacacatcc	780
ttttggatcca	atcctctccg	cagatacagg	cccttcgcta	ggagtcatag	ttttagattc	840
cataaggagg	agaaattgat	gaagatctgt	aagattaaaa	atctttcagg	ctcttcagaa	900
atagggcagg	gagcatgggt	caaagcgaag	atgctaacca	aagacaggag	actgtcaaa	960



aaaaagaaga	atactcaagg	actccagggt	ccaatcataa	cagtcaatgc	ctgctctcag	1020
agtgaccagt	tgaatccaga	gccaggagaa	aacagcatct	ctgaagagga	gtacagcaag	1080
aactgggttca	cagtggtccaa	atttagtcac	acagggtgtag	aacctttacat	acatcagaaa	1140
atgaaaaacta	aagaaattgg	acaatgtgct	atacaaatca	gtgattacct	aaagcagctct	1200
caagagtctg	ctcaggatct	cagcaaaaac	tctttgtgga	attccaggag	caccaacctc	1260
aataggaact	ccctgtcctc	tctaatttca	gagatctcag	cctccttaaa	aagccctcaa	1320
gagcctcacc	atcattattc	accttcactt	ctctttgcag	caggagaaga	aataactgtc	1380
tacaggttgg	aggagagtcc	ccctttaaac	cttgataaaa	gcatgtcctc	ttggtctcag	1440
cgtgggagag	cggcaatgat	ccagggtattg	tcccagagagg	agatggatgg	gggcctccgt	1500
aaagctatga	gagtcgtcag	cacttgggtct	gaggatgaca	ttctcaagcc	gggacaagtt	1560
ttcattgtca	agtcctttct	tcttgagggt	gtgcggacat	ggcataaaat	cttccaggag	1620
agcactgtgc	ttcatctttg	cctcagggaa	attcaacaac	aaagagctgc	tcaaaaattg	1680
atctatacct	tcaaccaagt	gaaaccacaa	accataccct	acacaccaag	gttcttgga	1740
gttttcttaa	tctactgcca	ttcagccaac	cagtggttga	ccattgagaa	gtatatgaca	1800
ggggagttcc	ggaagtataa	caacaacaat	ggtgatgaaa	tcacccccac	caacacctg	1860
gaggagctga	gtttggcttt	ctctcactgg	acctatgagt	acactcgggg	agagctgctg	1920
gttttagatt	tgcaagggtg	tggagaaaat	ttgacagatc	catctgttat	aaaacctgaa	1980
gtcaaacat	caagaggaat	ggtgtttgga	ccggccaatt	tgggggaaga	tgcaattaga	2040
aacttcattg	caaaacatca	ttggaactcc	tgctgccgga	agctcaaact	cccggattta	2100
aaaagaaatg	actattcccc	tgaaggata	aattccacct	ttggacttga	gataaaaata	2160
gaatcagctg	aggagcctcc	agcaaggagg	acgggtagaa	attccccaga	agatgatatg	2220
caactataaa	aaggaggagg	caagaagatc	ccagtgcttg	ccctgcctgc	caggaactct	2280
gtgataacat	agattgatca	acgtgatgtt	gattacatca	gcgtctcctt	gggacacgcc	2340
ttctgagcct	cacatctcct	tctgttcaaa	ggcctcattg	gtatatgatc	aatgggttct	2400
cctagacact	gacctctgtc	cagggcactt	tgcagctcca	tcctcaagtt	ccacacgaag	2460
atgcttggat	gagtcagctg	ggaatattgt	tcttgtgtac	ctcattgctt	tagctggtca	2520
cttggaaact	tggagcagaa	tcttcacat	taaaggatgg	ggttgggggg	gatacattta	2580
ttttattttc	tcactatgta	tgcagactng	acccccact	actatttgtc	acctcaccca	2640
cagattgtat	ttatgtctat	atatatgttt	cataaaaagc	ttaccaaac	caaaaaaaaa	2700
aaaaaaaaaa	aaaaaaaaaa	aa				2722

<210> 37

<211> 1924

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4787433CB1

<400> 37

cggcgccag	tgacggcgc	caccatgggg	tcccgccact	tcgaggggat	ttatgaccac	60
gtggggcact	tcggcagatt	ccagagagtc	ctctatttca	tatgtgcctt	ccagaacatc	120
tcttgtggta	ttcactactt	ggcttctgtg	ttcatgggag	tcacccctca	tcattgtctgc	180
aggccccag	gcaatgtgag	tcagggttgt	ttccataatc	actctaattg	gagtttgag	240
gacaccgggg	ccctgttgtc	ttcaggccag	aaagattatg	ttacggtgca	gttgacagaat	300
ggtgagatct	gggagctctc	aaggtgtagc	aggaataaga	gggagaacac	atcgagtttg	360
ggctatgaat	acactggcag	taagaaagag	tttcttctgt	tggatggcta	catatatgac	420
cagaacacat	ggaaaagcac	tgcggtgacc	cagtggaacc	tgggtctgtga	ccgaaaatgg	480
cttgaatgc	tgatccagcc	cctattttatg	tttggagtcc	tactgggatc	ggtgactttt	540
ggctactttt	ctgacaggct	aggacgcggg	gtgggtctgt	gggccacaag	cagtagcatg	600
ttttgttttg	gaatagcagc	ggcgtttgca	gttgattatt	acaccttcac	ggctgctcgc	660
ttttttcttg	ccatgggttg	aagtggctat	cttgtgtgtg	ggtttgtcta	tgtgatggaa	720
ttcattggca	tgaagtctcg	gacatgggag	tctgtccatt	tgcatctcct	ttttgcagtt	780
ggaaccttgc	tgttggtctt	gacaggatac	ttggtcagga	cctggtggct	ttaccagatg	840
atcctctcca	cagtgaactg	ccccctttatc	ctgtgctgtt	gggtgctccc	agagacacct	900
ttttggtctc	tctcagaggg	tctcagatgaa	gaagcacaaa	aaatagttga	catcatggcc	960
aagtggaaac	gggcaagctc	ctgtaaactg	tcagaacttt	tatcactgga	cctacaaggt	1020
cctgttagta	atagccccac	tgaagttcag	aagcacaccc	tatcatatct	gttttataac	1080
tggagcatta	cgaagaggac	acttacgggt	tggctaattct	ggttcactgg	aagtttgagg	1140
ttctactcgt	tttcttgtaa	ttctgttaac	ttaggaggca	atgaatactt	aaacctcttc	1200
ctcctgggtg	tagtggaat	tccgcctac	accttcgtgt	gcatcgccat	ggacaaggtc	1260
gggaggagaa	cagtcctggc	ctactctctt	tctgcagtg	cactggcctg	tgggtgctgt	1320

atggtgatcc	cccagaaaca	ttatatTTTg	ggtgtggtga	cagctatggt	tggaaaattt	1380
gccatcgggg	cagcatttgg	cctcatttat	ctttatacag	ctgagctgta	tccaaccatt	1440
gtaagatcgc	tggctgtggg	aagcggcagc	atggtgtgtc	gcctggccag	catcctggcg	1500
ccgttctctg	tggacctcag	cagcatttgg	atcttcatac	cacagttgtt	tgttgggact	1560
atggccctcc	tgagtggagt	gttaacacta	aagcttccag	aaacccttgg	gaaacggcta	1620
gcaactactt	gggaggaggc	tgcaaaactg	gagtcagaga	atgaaagcaa	gtcaagcaaa	1680
ttactttctc	caactaataa	tagtgggctg	gaaaaaacgg	aagcgattac	ccccagggat	1740
tctggtcttg	gtgaataaat	gtgccatgcc	tgtgtcttag	cacctgaaat	attatTTTacc	1800
ctaatgcctt	tgtattagag	gaatcttatt	ctcatctccc	atatgttgtt	tgtatgtctt	1860
tttaataaat	tttgaagaa	aatttttaaag	caaatatgtt	ataaaagaaa	taaaaactaa	1920
gatg						1924

<210> 38

<211> 1797

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7483598CB1

<400> 38

agcatccccg	tcccggagga	aaaaacattt	gtctggcgaa	ctccgggtgg	aaagcgcccc	60
aggctgccac	agcctagaga	tcttggggct	tcagcccctc	gcggcctgcc	gagggagcag	120
ggggcgcccc	tggaaactggc	tccctgcagc	tctgcggcta	cacgcggacc	tcggctgtgt	180
gcgaggtggc	ggaggaggct	ggccgggtgc	gaatccgtac	ccagccccag	catcttccac	240
ctgctgagga	ccaccgctca	gccatgggct	accagaggca	ggagcctgtc	atcccgccgc	300
agagagattt	agatgacaga	gaaacccttg	tttctgaaca	tgagtataaa	gagaaaacct	360
gtcagctctg	tgctcttttt	aatgtttgtc	actcgattat	aggatctggt	ataataggat	420
tgccttattc	aatgaagcaa	gctgggtttc	ctttgggaat	attgctttta	ttctgggttt	480
catatgttac	agacttttcc	cttgttttat	tgataaaagg	agggggccctc	tctggaacag	540
atacctacca	gtctttggtc	aataaaactt	tcggctttcc	agggtatctg	ctcctctctg	600
ttcttcagtt	tttgtatcct	tttatagcaa	tgataagtta	caatataata	gctggagata	660
ctttgagcaa	agtttttcaa	agaatcccag	gagttgatcc	tgaaaacgtg	tttatattgtc	720
gccacttcac	tattggactt	tccacagtta	cctttactct	gcctttatcc	ttgtaccgaa	780
atatagcaaa	gcttggaaaag	gtctccctca	tctctacagg	tttaacaact	ctgattcttg	840
gaattgtaat	ggcaagggca	atttcaactg	gtccacacat	acaaaaaaca	gaagacgctt	900
gggtatttgc	aaagcccaat	gccattcaag	cggtcggggg	tatgtctttt	gcattttattt	960
gccaccataa	ctccttctta	gtttacagtt	ctctagaaga	acccacagta	gctaagtggg	1020
cccgccattt	ccatatgtcc	atcgtagatt	ctgtatttat	ctgtatatte	tttgctacat	1080
gtggatactt	gacattttact	ggcttcaccc	aaggggactt	atttgaaaaat	tactgcagaa	1140
atgatgacct	ggtaacattt	ggaagatttt	gttatgggtg	cactgtcatt	ttgacatacc	1200
ctatggaatg	ctttgtgaca	agagaggtaa	ttgccaatgt	gttttttggg	gggaatcttt	1260
catcggtttt	ccacattgtt	gtaacagtga	tggatcatcac	tgtagccacg	cttgtgtcat	1320
tgtgtattga	ttgcctcggg	atagttctag	aactcaatgg	tgtgctctgt	gcaactcccc	1380
tcatttttat	cattccatca	gcctgttatc	tgaagaactg	tgaagaacca	aggacacact	1440
ccgataagat	tatgtcttgt	gtcatgcttc	ccattgggtg	tgtgggtgat	gttttttggg	1500
tcgtcatggc	tattacaaat	actcaagact	gcacccatgg	gcaggaaatg	ttctactgct	1560
ttcctgacaa	tttctctctc	acaaatacct	cagagtctca	tgttcagcag	acaacacaac	1620
tttctacttt	aaatattagt	atctttcaat	gagttgactg	ctttaaaaat	atgtatgttt	1680
tcatagactt	taaaacacat	aacattttacg	cttgcttttag	tctgtattta	tgttatataa	1740
aattattatt	ttggctttta	tcaagacttg	gcttttatga	gtagtgcaat	ataaaaaa	1797

<210> 39

<211> 3277

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7484823CB1

<400> 39

atgggggaag	gagatagtgg	gttgaggagc	ccaagaggcc	cacaggaagg	tgggaaatgc	60
agtagggaaa	aacagaagcg	caatatggag	gaactgaaga	aggaagtggg	catggatgat	120
cacaaattaa	ccttggaaga	gctgagcacc	aagtactccg	tggacctgac	aaagggccat	180
agccaccaaa	gggcaaaagg	aatcctgact	cgaggtggac	ccaatactgt	tacccacacc	240
cccaccactc	cagaatgggt	caaattctgt	aagcaactgt	tcggaggcct	ctccctccta	300
ctatggactg	gggccattct	ctgctttgtg	gcctacagca	tccagatata	tttcaatgag	360
gagcctacca	aagacaacct	ctacctgagc	atcgtaactgt	ccgtcgtggg	catcgctact	420
ggctgcttct	cctattatca	ggaggccaag	agctccaaga	tcatggagtc	ttttaagaac	480
atggtgcctc	agcaagctct	ggtaattcga	ggaggagaga	agatgcaaat	taatgtacaa	540
gaggtgggtg	tgggagacct	ggtggaaatc	aaggggtggag	accgagtcct	tgctgacctc	600
cggcttatct	ctgcacaagg	atgtaagggt	gacaactcat	ccttgactgg	ggagtcagaa	660
ccccagagcc	gctcccctga	cttcacccat	gagaaccctc	tggagaccgg	aaacatctgc	720
ttcttttcca	ccaactgtgt	ggaaggaacc	gcccggggta	ttgtgattgc	tacgggagac	780
tccacagtga	tgggcagaat	tgcctccctg	acgtcaggcc	tggcgggttg	ccagacacct	840
atcgctgctg	agatcgaaac	cttcacccat	ctgatcactg	tgggtggcgt	cttccttggg	900
gtcacttttt	ttgcgctctc	acttctcttg	ggctatgggt	ggctggaggc	tatcattttt	960
ctcattggca	tcattgtggc	caatgtgcct	gaggggctgt	tggccacagt	cactgtgtgc	1020
ctgacctca	cagccaagcg	catggcgcg	aagaactgcc	tgggtgaagaa	cctggaggcg	1080
gtggagagcg	tgggctccac	gtccaccatc	tgctcagaca	agacgggcac	cctcaccag	1140
aaccgcata	ccgtcgccca	catgtggttt	gatatgaccg	tgatgaggc	cgacaccact	1200
gaagaacaga	ctggaaaaac	atttaccaag	agctctgata	cctgggttat	gctggcccca	1260
atcgctggcc	tctgcaaccg	ggctgacttt	aaggctaate	aggagatcct	gcccattgct	1320
aagagggcca	caacaggtga	tgcttccgag	tcagccctcc	tcaagttcat	cgagcagctc	1380
tacagctctg	tggcggagat	gagagagaaa	aaccccaagg	tggcagaggt	tccttttaat	1440
tctaccaaca	agtaccagat	gtccatccac	cttcgggagg	acagctccca	gacccacgta	1500
ctgatgatga	agggtgctcc	ggagaggatc	ttggagtgtt	gttctacctt	tcttctgaat	1560
gggcaggagt	actcaatgaa	cgatgaaatg	aaggaagcct	tccaaaatgc	ctatttagaa	1620
ctgggagggtc	tgggggaacg	tgtgctaggg	ttctgcttct	tgaatctgcc	tagcagcttc	1680
tccaagggat	tcccatttaa	tacagatgaa	ataaatttcc	ccatggacaa	cctttgtttt	1740
gtgggcctca	tatccatgat	tgacctccc	cgagctgcag	tgctgatgac	tgtgagcaag	1800
tgtcgcatga	cagggaattaa	ggtgatcatg	gtaacaggag	atcatcccat	tacagctaag	1860
gccattgccca	agggtgtggg	catcatctca	gaaggcactg	agacggcaga	ggaagtgcgt	1920
gcccggctta	agatccctat	cagcaaggtc	gatgccagtg	ctgccaaaagc	cattgtgggtg	1980
catggtgcag	aactgaagga	catacagtc	aagcagcttg	atcagatcct	ccagaaccac	2040
cctgagatcg	tggttgctcg	gacctcccct	cagcagaagc	tcattcattgt	cgagggatgt	2100
cagaggctgg	gagccgttgt	ggccgtgaca	ggtgacgggg	tgaacgactc	ccctgcgctg	2160
aagaaggctg	acattggcat	tgccatgggc	atctctggct	ctgacgtctc	taagcaggca	2220
gccgacatga	tctgtctgga	tgacaacttt	gcctccatcg	tcacgggggt	ggaggagggc	2280
cgctgatct	ttgacaacct	gaagaaatcc	atcatgtaca	ccctgaccag	caacatcccc	2340
gagatcacgc	ccttcctgat	gttcatcatc	ctcgggtatac	ccctgcctct	gggaaccata	2400
accatcctct	gcattgatct	cggcactgac	atggtccctg	ccatctcctt	ggcttatgag	2460
tcagctgaaa	gcgacatcat	gaagaggctt	ccaaggaacc	caaagacgga	taatctgggtg	2520
aaccaccgtc	tcattggcat	ggcctatgga	cagattggga	tgatccaggc	tctggctgga	2580
ttctttacct	actttgtaat	cctggctgag	aatgggttta	ggcctgttga	tctgctgggc	2640
atccgcctcc	actgggaaga	taaatacttg	aatgacctgg	aggacagcta	cggacagcag	2700
tggacctatg	agcaacgaaa	agttgtggag	ttcacatgcc	aaacggcctt	ttttgtcacc	2760
atcggtggtg	tgcatgtggc	ggatctcatc	atctccaaga	ctcgccgcaa	ctcacttttc	2820
cagcagggca	tgagaacaaa	agtcttaata	tttgggatcc	tggaggagac	actcttggct	2880
gcatttctgt	cctacactcc	aggcatggac	gtggccctgc	gaatgtaccc	actcaagata	2940
acctgggtggc	tctgtgccat	tccctacagt	attctcatct	tcgtctatga	tgaaatcaga	3000
aaactcctca	tccgtcagca	cccggatggc	tgggtggaaa	gggagacgta	ctactaaact	3060
cagcagatga	agagcttcat	gtgacacagg	ggtgttgtga	gagctgggat	ggggccagag	3120
attataagtt	tgacacaaca	tctgagacac	taggatgaat	tatcttggat	gagaaagatg	3180
ggcaatcctg	ggctggcttg	agggaaatcat	gggcagagga	tgaggtgggc	tgaagggaa	3240
cccagcctgc	atctagctgg	agccccgcag	ggagggg			3277

<210> 40  
 <211> 2773  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> misc\_feature

<223> Incyte ID No: 143935CB1

<400> 40

```
atggaagaga tgtcaggaga aagtgtggtg agctcagcgg tgccagcggc tgctaccgc 60
accaattcct tcaagggcac gagccccagc tccaaatagc tgaagctgaa tgtgggtgga 120
gccctctact ataccacat gcagacgctg accaagcagg acaccatgct gaaggccatg 180
ttcagcgggc gcatggaagt gctcaccgac agtgaaggct ggatcctcat tgaccgctgt 240
gggaagcact ttggtacgat actcaactac cttcgagacg gggcgggtgcc tttaccgcag 300
agccgcccgg agatcgagga gctgctagca gaagccaagt actacctagt ccaaggcctg 360
gtggaagagt gccaggcggc cctacaaaac aaagatactt atgagccttt ctgcaaggtc 420
cctgtgatca cctcatccaa ggaagaacaa aaacttatag cgacttcaaa taagccagcc 480
gtgaagttgc tctacaacag aagtaacaac aaatactcat ataccagcaa ttctgacgac 540
aatatgttga aaaacattga actgtttgat aagctgtctc tgcgctttaa cggaagggtc 600
ctgttcataa aggatgttat tggggatgaa atctgctgct ggtcctttaa tggtcagggc 660
cggaagattg ctgaagtctg ttgtacctcc atcgtctatg ccaactgagaa gaaacagacc 720
aaggtggagt ttcccgaagc ccggatttat gaggagaccc tgaacatttt gctgtatgag 780
gcccaggatg gccggggacc tgacaatgcg ctctggagg ccacaggcgg ggcggcggg 840
cgctcccacc acctggacga ggacgaggag cgggagcggg tcgagcgcgt gcggaggatc 900
cacatcaagc gccctgatga ccgggcccac ctccaccagt gagcaggcaa gagaccgagc 960
cgccctctct tcaccgcccc cactccctgc cgtgctacac ccagatcctg tgcaggctgc 1020
cgggcccctt ctgcttccct tggagcctgg agatactttt gtaacaagcc agatgattat 1080
tttggatttg cttgacaagg caaattgatt gtcttgacce aggcgtatga cccctgtcgt 1140
tgaacaagct gtgtctaaga tctctacttt tcatgagaat ctgagactct ttggagccag 1200
gctttctcgg ttctcagagg aaaagtatga atgagtgtga agtgtatgtg agaacttttg 1260
tttgaatat ttatttttgt ggggtgtcgac tctctatgtg ggcttttttg gtgacactcc 1320
cttaagggtt cagtttgaca attctgagag ttgtcctgca gttggaggcc accagaggta 1380
tctgagctcc ctgcttccca tttcataatc ctccagcccc agcaggteca ctctgggttc 1440
ctgtgtgttt ggcccgggca caatccccac tgctttgcta gacgtgcttt ctgccatgtg 1500
gctttggggc tagagcttgt tgataattgc agcttgtggc agtggaaata tggctgaatg 1560
agcgtctaaa tcggtgagac cagtgcacac ttgggtgcaa ggctttgttt agggatcaag 1620
ccttttgcca ccttgggctg gtctttggcc tgggtgtcac tgggacccca tatgtctgcy 1680
taggagcaga actttccatg gcagtaagtg tccagctctg tttctgggtc tttccccaac 1740
tccagccccg tccagttggt ctctgtattg acccgactcc actccaggaa ggccatctga 1800
ccctgtgaca ggcatagtct ataaactacc cctccctggg atcccgtccc tcttcagcct 1860
ccttccccat gaagctgggc taactttcta agtcattttg cttagaaatt cagtgtggcc 1920
catacccttt gtccctccag cctggcatcc aggcagggac accctcacac caccagcccc 1980
agggagcttc cctgctataa acacagaccc ccttgtcttt gcctctgatt tttacacagt 2040
gtagagtggc cagcagtga caggttgagg atgtgcgggt agatagataa ctttgggtct 2100
ggtttgtgtc tgtgttcatt tttgtttaag ggatatgtgt gactgtgggt ggggacgtgt 2160
gcttgtgggg cacagggtggc gggccctgct ggagcccggc tgggcccagc gcctatgtag 2220
gacgggtggt ctcagtgacc tacctcccag gctcctctgc acctgcaaag gaacaggagt 2280
gagtcgtgac tgacaggggt ggttgagact agactaggta gagtagttac caggagatgt 2340
gaatgtgcgt caggtgatgg atgggtttgt caagggaatc gttaccgttt tataccaaag 2400
gtattaacat gggcagcctt tgacacatgt attccaaaaa cgagtttata ttttcaaacy 2460
gtttttacag cttagacttt gtaacttact ccctgctgtg gacagttgta tgccttcatt 2520
ttgtatccaa cagcaaaagtc tacaataaaa ctttaaaaca atcatgactg aatgtcaaaa 2580
tcgtgtattg ggcagatgct ttttaaaact tcgtgtgaga aacttttata ttagggcatt 2640
tggattttat taagtgttaa ggaagagggg cttacaaaat gtttcgtaaa tattttatac 2700
tgtttaagtg ttaaacacca accctgtctt tcttttgggt tgagcttttt tagaaagtcg 2760
aagtgaatgt tgg 2773
```

<210> 41

<211> 2966

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5923789CB1

<400> 41

```
tttttccaag tgagagactg ccccatatgt tagtgcaata tgtcactgga ggtgaagcat 60
cagttgtatt ggtgggaacc tgccgtttgc tgccccctt ttctcatgc cttttctctg 120
```

ctctctgac	ttttctaggt	ctctggccta	tcaggaggac	aactgggtgct	gcaatagaag	180
ccagtggcta	agtctctgt	atggcggtgt	taagggttga	gcctctcacc	tctgccttcc	240
tccatttttg	gctgggtacc	tttgtgtct	tcctgaatgg	tcttcgagca	gaggctggtg	300
gctcagggga	cgtgccaaag	acagggcaga	acaatgagtc	ctgttcaggg	tcctcggact	360
gcaaggagg	tgctacccct	ccaatctgg	acccggagaa	cccttccctt	ggggacaaga	420
ttgccagggt	cattgtctat	tttgtggccc	tgatatacat	gttccttggg	gtgtccatca	480
ttgctgaccg	cttcatggca	tctattgaag	tcacacctc	tcaagagagg	gagggtgaca	540
ttaagaaacc	caatggagaa	accagcacaa	ccactattcg	ggctctggaat	gaaactgtct	600
ccaacctgac	ccttatggcc	ctgggttcct	ctgctcctga	gatactcctc	tctttaattg	660
agggtgtgtg	tcattgggttc	attgtctgtg	atctgggacc	ttctaccatt	gtaggagtg	720
cagccttcaa	catgttcac	atcattggca	tctgtgtcta	cgtgatccca	gacggagaga	780
ctcgcaagat	caagcatcta	cgagctctct	tcacacccg	tgcttggagt	atctttgcct	840
acatctggct	ctatatgatt	ctggcagctt	tctcccttgg	tggtgtccag	gtttgggaag	900
gcctcctcac	tctcttcttc	tttccagtgt	gtgtccttct	ggcctgggtg	gcagataaac	960
gactgctctt	ctacaaatc	atgcacaaa	agtaccgcac	agacaaacac	cgaggaatta	1020
tcatagagac	agagggtgac	caccctaagg	gcattgagat	ggatgggaaa	atgatgaatt	1080
cccattttct	agatgggaac	ctgggtgccc	tggaaggga	ggaagtggat	gagtcccgca	1140
gagagatgat	ccggattctc	aaggatctga	agcaaaaaca	cccagagaag	gacttagatc	1200
agctgggtga	gatggccaat	tactatgtct	tttcccacca	acagaagagc	cgtgccttct	1260
accgtatcca	agccactcgt	atgatgactg	gtgcaggcaa	tatcctgaag	aaacatgcag	1320
cagaacaagc	caagaaggcc	tccagcatga	gcgagggtga	caccgatgag	cctgaggact	1380
ttattttcaa	ggtcttcttt	gacccatgtt	cttaccagt	cctggagaac	tggtgggctg	1440
tactcctgac	agtgggtgag	aaagggggag	acatgtcaaa	gaccatgtat	gtggactaca	1500
aaacagagga	tggttctgcc	aatgcagggg	ctgactatga	gttcacagag	ggcacggtg	1560
ttctgaagcc	aggagagacc	cagaaggagt	tctccgtggg	cataattgat	gacgacattt	1620
ttgaggagga	tgaacacttc	tttgaagg	tgagcaatgt	cgcgatagag	gaggagcagc	1680
cagaggagg	gatgcctcca	gcaatattca	acagtcttcc	cttgccctcg	gctgtcctag	1740
cctccccttg	tgtggccaca	gttaccatct	tggtatgatga	ccatgcaggc	atcttcaact	1800
ttgaatgtga	tactattcat	gtcagtga	gtattggtgt	tatggaggtc	aagggtctgc	1860
ggacatcagg	tgcccggggt	acagtcatcg	tcccctttag	gacagtagaa	gggacagcca	1920
agggtggcgg	tgaggacttt	gaagacacat	atggggagtt	ggaattcaag	aatgatgaaa	1980
ctgtgaaaac	cataagggtt	aaaatagtag	atgaggagga	atacgaagg	caagagaatt	2040
tcttcattgc	ccttggtgaa	ccgaaatgga	tggaacgtgg	aatatcagat	gtgacagaca	2100
ggaagctgac	tatggaagaa	gaggaggcca	agaggatagc	agagatggga	aagccagtat	2160
tggttgaaca	ccccaaacta	gaagtcatca	ttgaagagtc	ctatgagttc	aagactacgg	2220
tggacaaact	gatcaagaag	acaaacctgg	ccttggttgt	ggggacccat	tcctggagg	2280
accagttcat	ggaggccatc	accgtcagtg	cagcagggga	tgaggatgag	gatgaatccg	2340
gggaggagag	gctgccctcc	tgctttgact	acgtcatgca	cttcctgact	gtcttctgga	2400
agggtgctgt	tgctgtgtg	ccccccacag	agtactgcca	cggctgggccc	tgcttcgccc	2460
tctccatcct	catcattggc	atgctcaccg	ccatcattgg	ggacctggcc	tcgcacttcg	2520
gctgcaccat	tggtctcaaa	gattcagtca	cagctgttgt	tttcgtggca	tttggcacct	2580
ctgtcccaga	tacgtttgcc	agcaaagctg	ctgccctcca	ggatgtatat	gcagacgcct	2640
ccattggcaa	cgtgacgggc	agcaacgcgc	tcaatgtctt	cctgggcac	ggcctggcct	2700
gggtccgtgg	cgccatctac	tggtctctgc	agggacagga	gttccacgtg	tcggccggca	2760
cactggcctt	ctccgtcacc	ctcttcacca	tctttgcatt	tgtctgcac	agcgtgctct	2820
tgtaccgaag	gcggccgcac	ctgggagggg	agcttgggtg	ccccctggc	tgcaagctcg	2880
ccacaacatg	gctctttgtg	agcctgtggc	tctctacat	actctttgcc	acactagagg	2940
cctattgcta	catcaagggg	ttctaa				2966

<210> 42

<211> 4404

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6046484CB1

<400> 42

atgtatatcc	gtgtatecta	tgacaccaag	ccagactcac	tgctccatct	catggtgaaa	60
gattggcagc	tggaaactcc	caagctctta	atatctgtgc	atggaggcct	ccagaacttt	120
gagatgcagc	ccaagctgaa	acaagctctt	gggaaaggcc	tgatcaaggc	tgctatgacc	180
accggggcct	ggatcttcac	cggggggtgc	agcacagggt	ttatcagcca	cgtaggggat	240

gccttgaaag	accactcctc	caagtcacaga	ggccggggtt	gtgctatagg	aattgctcca	300
tggggcatcg	tggagaataa	ggaagacctg	gttggaagg	atgtaacaag	agtgtaccag	360
accatgtcca	accctctaag	taagctctct	gtgctcaaca	actccacac	ccacttcac	420
ctggctgaca	atggcacctc	gggcaagtat	ggcgccgagg	tgaacttcg	aagacaactg	480
gaaaagcata	tttcaactca	gaagataaac	acaagaatcg	gtcaagggtg	tcctgtggtg	540
gcactcatag	tggaaggagg	acccaatgtg	atctcgattg	ttttggagta	ccttcgagac	600
accctcccg	tgccagtgg	tgtctgtgat	gggagtggac	gggcatcgga	catcctggcc	660
tttgggcata	aatactcaga	agaaggcgga	ctgataaatg	aatctttgag	ggaccagctg	720
ttggtgacta	tacagaagac	tttcacatac	actcgaaccc	aagctcagca	tctgttcac	780
atcctcatgg	agtgcataaa	gaagaaggaa	ttgattacgg	tatttcggat	gggacagaa	840
ggacaccagg	acattgattt	ggctatcctg	acagctttac	tcaaaggagc	caatgcctcg	900
gccccagacc	aactgagctt	agcttttagcc	tggaaacagag	tcgacatcgc	tcgcagccag	960
atctttatatt	acgggcaaca	gtggccggtg	ggatctctgg	agcaagccat	gttggatgcc	1020
ttagttcttg	acagagtggg	ttttgtgaaa	ttactcatag	agaatggagt	aagcatgcac	1080
cgttttctca	ccatctccag	actagaggaa	ttgtacaata	cgagacatgg	gccctcaaat	1140
acattgtacc	acttggtcag	ggatgtcaaa	aaggggaaac	tgcccccaga	ctacagaatc	1200
agcctgattg	acatcgccct	ggtgatcgag	tacctgatgg	gcggggctta	tcgctgcaac	1260
tacacgcgca	agcgcttcgg	gacctctac	cacaacctct	tcggcccaa	gagggatgat	1320
attcccttga	ggcgaggaag	aaagacaacc	aagaaacgtg	aagaagaggt	ggacattgac	1380
ttggatgac	ctgagatcaa	ccacttcccc	ttccctttcc	atgagctcat	ggtgtgggct	1440
gttctcatga	agcggcagaa	gatggccctg	ttcttctggc	agcacgggtg	ggaggccatg	1500
gccaaggccc	tgggtggcctg	caagctctgc	aaagccatgg	ctcatgaggc	ctctgagaac	1560
gacatgggtg	acgacatttc	ccaggagctg	aatcacaatt	ccagagactt	tggccagctg	1620
gctgtggagc	tcctggacca	gtcctacaag	caggacgaac	agctggccat	gaaactgctg	1680
acgtatgagc	tgaagaactg	gagcaacgcc	acgtgcctgc	agcttgccgt	ggctgccaaa	1740
caccgcgact	tcacgcgcga	cacgtgcagc	cagatgctgc	tcaccgacat	gtggatgggc	1800
cggtccgcga	tgcgcaagaa	ctcaggccctc	aaggtaattc	tgggaattct	tctaccccc	1860
accatcttgt	ttttggaatt	tcgcacatat	gatgatttct	cgtatcaaac	atccaaggaa	1920
aatgaggatg	gcaaagaaaa	agaagaggaa	aatacggatg	caaatgcaga	tgctggctca	1980
agaaaggggg	atgaggagaa	cgagcacaaa	aaacagagaa	gtattcccat	cggaaacaaag	2040
atctgtgaat	tcataaacgc	gccattgtgc	aagtctctgt	tctttcagat	atcatacttg	2100
ggctacctgc	tgtctgttaa	ctacgtcatc	ctgggtcgga	tggatggctg	gccgtccctc	2160
caggagtggg	tcgtcatctc	ctacatcgtg	agcctggcgt	tagagaagat	acgagagggtg	2220
gtactccaa	aggcctcccc	cagcccacta	gccagaaaat	caaagtcttg	cttccaggag	2280
tactggaaca	tcacagatct	cgtggccatt	tcacattca	tgattggagc	aattcttcgc	2340
ctacagaacc	agccctacat	gggctatggc	cgggtgatct	actgtgtgga	tatcatcttc	2400
tggtagatcc	gtgtcctgga	catctttggt	gtcaacaagt	atctggggcc	atacgtgatg	2460
atgattggaa	agatgggttag	tagtggcatc	ctgtgggtag	tcacatgct	ggtcgtgctc	2520
atgagtttcg	gagtagcccg	tcaagccatt	ctgcatccag	aggagaagcc	ctcttggaag	2580
ctggcccgaa	acatcttcta	catgccttac	tggatgatct	atggagaggt	gtttgcagac	2640
cagatagacc	gtaagagttt	ttttctttca	gctccttggt	gtgagaacct	atatgatgag	2700
gagggcaagc	ggcttctctc	ctgtatcccc	ggcgccctgc	tcactccagc	actcatggcg	2760
tgtatcttac	tggtcgcaaa	catcctgctg	gtgaacctgc	tgattgctgt	gttcagcaat	2820
accttctttg	aagtaaaatc	aatatccaac	caggtgtgga	agttccagcg	atatcagctg	2880
attatgacat	ttcatgacag	gccagtcctg	ccccaccgga	tgatcatttt	aagccacatc	2940
tacatcatca	ttatgctct	cagcgccgcg	tcgaggaaaa	agagagaagg	ggaccaagag	3000
gaacgggagc	gtggattgag	tatgttcctt	agcgacgagg	agctaaagag	gctgcatgag	3060
ttcgaggagc	agtgcgtgca	ggagcacttc	cgggagaagg	aggatgagca	gcagtcgtcc	3120
agcgacgagc	gcatecggtt	cacttctgaa	aggggtgaaa	atatgtcaat	gaggttggaa	3180
gaaatcaatg	aaagagaaac	ttttatgaaa	acttccctgc	agactgttga	ccttcgactt	3240
gctcagctag	aagaattatc	taacagaatg	gtgaatgctc	ttgaaaatct	tgcggaatc	3300
gacaggtctg	acctgatcca	ggcaggtccc	cgggcttctt	ctgaatgtga	ggcaacgtat	3360
cttctccggc	aaagcagcat	caatagcgct	gatggctaca	gcttgatcgc	atatcatttt	3420
aacggagaag	agttattatt	tgaggataca	tctctctcca	cgtcaccagg	gacaggagtc	3480
aggaaaaaaa	cctgttcctt	ccgtataaag	gaagagaagg	acgtgaaaac	gcacctagtc	3540
ccagaatgtc	agaacagctc	tcacctttca	ctgggcacaa	gcacatcagc	aacccagat	3600
ggcagtcacc	ttgcttaag	aacgtgaaag	aacgtgaaag	agtcaaaatt	aggteccagat	3660
attgggattt	caaaggaaga	tgatgaaaga	cagacagact	ctaaaaaaga	agaaactatt	3720
ttcccaagtt	taaataaaac	agatgtgata	catggacagg	acaaatcaga	tgttcaaaac	3780
actcagctaa	cagtggaaac	gacaaatata	gaaggcacta	tttctatcc	cctggaagaa	3840
acaaaaatta	cacgctattt	ccccgatgaa	acgatcaatg	cttgtaaaac	aatgaagtcc	3900
agaagcttcg	tctattcccg	gggaagaaag	ctggctcggtg	gggttaacca	ggatgtagag	3960
tacagttcaa	tcacggacca	gcaattgacg	acggaatggc	aatgccaagt	tcaaaagatc	4020

acgcgctctc	atagcacaga	tattccttac	attgtgtcgg	aagctgcagt	gcaagctgag	4080
cataaagagc	agtttgcaga	tatgcaagat	gaacaccatg	tcgctgaagc	aattcctcga	4140
atccctcgct	tgtccctaac	cattactgac	agaaatggga	tggaaaactt	actgtctgtg	4200
aagccagatc	aaactttggg	attcccatct	ctcaggtcaa	aaagtttaca	tggacatcct	4260
aggaatgtga	aatccattca	gggaaagtta	gacagatctg	gacatgccag	tagtgtaagc	4320
agcttagtaa	ttgtgtctgg	aatgacagca	gaagaaaaaa	aggttaagaa	agagaaagct	4380
tccacagaaa	ctgaatgcta	gtct				4404

<210> 43  
 <211> 669  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7481427CB1

<400> 43						
atgacggaac	aagccatctc	cttcgccaag	gacttcctag	ctggaggcat	caccgccgct	60
atctccaaga	cagccgtggc	atcgatcaag	cgggtccagc	tgctgctgca	gatgcagcac	120
gccagcatgc	cgatggccgc	cgccaagcag	tgcaagggca	tcgtggactg	catcgtccgc	180
atccccaaag	atcagggcgt	gctgtccttc	tggaggggca	acctggccaa	tgtcatccgc	240
tactcccca	cgcaagccct	caatttcgcc	ttcaaggata	agtacaagca	gatcttctcg	300
gcgggcggtg	acaagcacac	gcagttctgc	aggtactttg	cgggcaacct	ggcctctggc	360
ggcacggccg	tcgtgtatcc	cctggatttc	accagaacct	gcctggcagc	cgacgtggga	420
aagtcaggca	cggagcgcga	gttccgaggg	ctgggagact	gcctagtga	gatcagcaag	480
tccgacggca	tccggggcct	ttaccagggc	ttcagcgtct	ccgtgcaggc	catcatcatc	540
taccagggcag	cctacttcag	ggtgtacgat	acggccaatg	gcatgttccc	cgaccccaag	600
aacacacaca	tcctggtgag	ctggatgacc	gcgcagaccg	tgacggccgt	ggctggcggtg	660
ctctcctaa						669

<210> 44  
 <211> 1823  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7483595CB1

<400> 44						
aaaaaaaaac	aaataaaaaa	tgcccagtg	gagcatctgt	agtcccagct	atttgggaag	60
ttgaagcaga	gaatcccttg	tgcccaggag	tttgagggtta	cagtgagcta	tgattgcacc	120
actgcactcc	ggcctgggca	acagagtgag	atcctgtctc	ataaataaat	aaataggcaa	180
aggcttggtc	tgatcacctt	gtccctgttt	tccccaagtc	aacctagaga	gcagtagatt	240
ctagggcaag	aacttaagtg	agctgactgc	tcacaaagaa	ggaaaggagt	agtgtgagct	300
agcacttggt	gggagcctct	gatcaacaat	acaaggacct	gtgttatcaa	cttctgttca	360
ttagggaaag	agagttacaa	gagttaccaa	aggggcctat	cttaactttt	ctatggaaac	420
ccttcttttc	ctggagagtg	ccattgggtc	cattgctgga	ttgaagacat	tctacccac	480
tctggttttc	tcttctctgc	agctaataat	gggcgtattg	ggtttgggct	tcattgccac	540
ttaccttccg	gagtcctgca	tgagtgtcta	cctggctgct	gtggcacttc	atatcatgct	600
gtcccagctg	actttcatct	ttgggattat	gattagtttc	catgccggtc	ccatctcctt	660
cttctatgac	ataattaatt	actgtgtagc	tctcccaaaa	gcgaattcca	ccagcattct	720
agtattttcta	actgttggtg	ttgctctgcg	aatcaacaaa	tgtatcagaa	tttctttcaa	780
tcagtatccc	attgagtttc	ccatggaatt	atttctgata	attttacaag	ccttctcctt	840
atctttgggtg	agctcccttc	tgctcatatt	tctgggcaag	aagattgccca	gtcttcacaa	900
ttacagtgtc	aattccaacc	aggatttaat	agccatcggc	ctttgcaatg	tcgtcagttc	960
atttttcaga	tcttggtgtg	ttactgggtc	tattgctagg	actattatcc	aggataaatc	1020
tggaggaagc	acaacagttt	gcactctctg	taggcgcaga	gctaagattc	ttctcctggg	1080
tcaaatccct	aacaccaaca	tttatagaag	catcaatgat	tatcgggaga	tcacacccat	1140
tcctgggggtg	aaaatcttcc	agtgtctcag	ctcaattaca	tttgtaaatg	tttactacct	1200
aaagcataag	ctgttaaaag	aggttgatat	ggtaaagggtg	cctcttaaaag	aagaagaaat	1260
tttcagcttg	tttaattcaa	gtgacaccaa	tctacaagga	ggaaagattt	gcaggtgttt	1320

ctgcaactgt	gatgatctgg	agccgctgcc	caggattctt	tacacagagc	gatttgaaaa	1380
taaactggat	cccgaagcat	cctccgttaa	cctgattcac	tgctcacatt	ttgagagcat	1440
gaacacaagc	caaaactgcat	ccgaagacca	agtgccatac	acagtatcgt	ccgtgtctca	1500
gaaaaatcaa	gggcaacagt	atgaggaggt	ggaggaagtt	tggttctcta	ataactcatc	1560
aagaaacagc	tcaccaggac	tgctgatgt	ggcggaagc	caggggagga	gatcactcat	1620
cccttactca	gatgcgtctc	tactgccag	tgtccacacc	atcatcctgg	atttctccat	1680
ggtacactac	gtggattcac	gggggttagt	cgtattaaga	caggtaaagta	ctgaggaggc	1740
cttggcagga	gctctgatcc	ccctcctacc	atcccaaccc	caccctgac	ctgattgaat	1800
tctgttagtc	actgatcagc	tgc				1823

<210> 45

<211> 2931

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3788427CB1

<400> 45

atgtccatag	ttcgtttatc	cgtccatgcc	aaatggatta	tggggaaggt	gactgggaca	60
aaaatgcaaa	agactgctaa	agtgaagtg	atcaggcttg	ttctggatcc	ccatttatta	120
aagtattata	ataagcaaaa	aacgtacttt	gtcaccaatg	cccttcagca	gtgcacaatt	180
ggggatatcg	tgcttctcaa	agctttacct	gttccacgaa	caaagcatgt	gaaacatgaa	240
ctggctgaga	ttgttttcaa	agttagaaaa	ctcgtagatc	ctgtgacagg	aaaacctcgt	300
gcaggaaacta	cctatctgga	gagtccattg	agttcgaaaa	ccaccaggg	cgtagacggg	360
gcttcccgtc	cctcacgcgg	gcccgcacca	tgccgggcgg	ggcctggggc	ccgcgcctc	420
cgccccgtggc	cggagagccc	caggcccag	ccacgcggtc	ttccggggcc	gggtcggggc	480
tcaatggcta	cgtggaggcg	ggacggccga	ctgacaggcg	gccaaggct	gctgtgcgct	540
gggctggcgg	ggacgctcag	cctcagcctc	accgcgcccc	tgagctcgc	caccgtgctg	600
gcccagggtg	gcgtcgtgcg	aggccacgcc	cggggacgct	gggccacagg	gcaccgggtg	660
tgccgggcag	aggggctccg	ggccctgtgg	aaggggaacg	cgggtggcgtg	cctgcgcctc	720
ttccccctgca	gcgcctgca	gctcgccgcc	taccgcaa	ttgttgtgct	gttcacagat	780
gacctggggc	acatttccca	tgaggctcc	atcatggctg	ggagtctcgc	aggcatgggt	840
tccaccattg	taacatatcc	tacagacctc	atcaaaaacc	ggttgatcat	gcagaacata	900
ctggaaccat	cgtacagggg	gctcctccat	gctttttcta	ctatttacca	acaggaaggg	960
ttccttgccc	tttatcgagg	ggtttccctc	actgtttag	gtgctctccc	gttctctgct	1020
ggctcccttc	ttgtttacat	gaacctggag	aaaatctgga	acggaccccg	agatcagttc	1080
tctctcccac	agaactttgc	taatgtctgt	ctggctgctg	cagtgaacca	gacctctctc	1140
tttccctttg	agacgtgaa	gagaaagatg	caggctcaga	gccccctac	ccacacagt	1200
ggaggagtag	atgtccattt	ctcaggagca	gtggactgct	tccggcagat	agtgaaggcc	1260
caggggggtcc	tggggctctg	gaatggattg	acagccaatt	tactgaagat	agttccatat	1320
tttggaatta	tgtttagcac	ctttgagttc	tgcaagagaa	tctgtcttta	tcaaaatgg	1380
tacattctgt	ctccactgag	ctataaattg	accccaggag	tcgatcagag	tttgcagccc	1440
caggaattac	gagaattaaa	gaagtctctc	aaaacgagaa	aactgaagcc	taaaaaacca	1500
actctataaa	atggaatgga	actagaagtg	cactgactga	cagcatgttg	caatcacaga	1560
taaagttagc	ctcataactg	tgacactgag	gagggatgag	aagctactgc	tatctgctgc	1620
tctgggaaat	gagatgggtt	ggccatgcac	cacctcagat	ctccaaactg	atgtccctgt	1680
gaacacctgc	actatgagaa	ggtttcccaa	caagagtgt	tcgatgatgg	attccacgag	1740
attttataac	cagagtctag	cagtataaaa	atgtacaatt	tacgctctga	atagaaacat	1800
ggcaccaaag	aaagggcatt	gccatgacat	ctcaaagcaa	caggtaattg	aagtggctga	1860
gtggatacca	cgttgtgagg	tcagccactg	ggtcacttgt	agtcaatgca	aagcttgac	1920
aaaccaacgt	cagcctagtt	tactcagtt	tgacctgaga	ctttccaata	caatctgttg	1980
ttttgttttg	tttttagaga	caggggtctt	ctctgtcacc	cagactggag	tacaatgaca	2040
caatcatagc	tcactgcagc	ctttaactac	tgggctcaag	acatcctcct	gcttcagcct	2100
ccagagagtt	gggaccatag	gtgcacacca	ccacacctag	ctaatttttg	ggggagggtc	2160
tgctatgttg	cccaggctgg	tcttgaactc	ctgggctcat	gcaatcctcc	tgcttggcc	2220
tcccaaagcg	ctaggattag	aggtgtgagc	cgctgcaccc	tgccccagta	caatcttttt	2280
tgaactcaaa	tttttgctga	catctgagtg	cacacaccac	agtgtaaatt	atgccttacc	2340
agaatctaaa	tgaaaatagc	gaacatttaa	aagctatcac	cattgtagta	gaatcatcct	2400
tcttttttga	aatttgaagc	atcccaggct	taaaatcttg	tgtttcagaa	agacagttta	2460
taccattgat	gcttaattg	cccccaaaag	accttctgat	tgaagtcatg	tacagttcag	2520
tggcctaaat	tctctgcctt	tttaacttgc	tttgcaagcc	tactctgaaa	ataagttatt	2580



tagtcaagtt	atttctcaaag	atgtcccagt	tgcctagaaa	ggatcaaatg	gaacatttga	2640
cacacatact	caaaaaaatg	taactgacta	taaacacttt	aacctaatca	tctgtatcaa	2700
acttttctaaa	aatcaaatct	caggattgtt	ccactttaga	gattctatgt	aaagtttata	2760
taactatact	tgtcaaatag	cacctatcta	tgcattttaa	aatgcattaa	tatctaacat	2820
aaaattgtta	atttatacac	acacatacac	ttatacatag	agagagattt	tcggtatgtt	2880
gcaagtattt	tttccgatta	aatttatatt	ccaaagaaaa	aaaaaaaaaa	a	2931

<210> 46  
 <211> 1492  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6972455CB1

<400> 46	
cttgcctcctt	gatcagtttg
tggaattcat	aaaaagctca
atagctggct	gcaagccacg
gggtgcgggaa	acctgaccac
gccttgacaga	cgatccgaac
tccttcgcct	ttttcggaga
tgtttgccag	cctgatgctg
tcgccctggg	caccctgaag
gccctgacat	gcaaccccg
tcaaggtcaa	gctggaactg
ctggcaaaagt	cgacctgggt
tcgacttctc	caacgcctac
ccatcgccag	cctcgacgac
aagacatcga	gctgacgcg
acaactccac	catcgccgcc
tggtcatggt	ggccatcagc
tcaaggattc	gccggtgtac
tcaaccagat	cctggtggca
ggctcaaaga	accgctgcct
cgactttgtg	ccggtgctgg
tgagctgacg	gccattggcg
acgggcgtgg	aagatccaac
caacacgcgg	ttcctgggtg
gaaaattacc	gagtggaag
caccgagatc	atccgtgccg
gcatccaggc	catcccacgg
gggcaactgg	ag
	60
gcttgggcaa	60
aaaactattc	120
gctcgaaaag	180
acaaacatca	240
ctcagaaaat	300
ctcactgcc	360
gatgacatcg	420
ggttcggtcg	480
gccgaccaac	540
ttcctcacca	600
gcaaagggtca	660
cctgatgccg	720
ggcgccatcg	780
tttgaagaca	840
agtggcaacg	900
aagggtgaaac	960
ctggaaaagg	1020
gcgatgcagt	1080
cctatcaatt	1140
tgttcaccct	1200
gcgccgtgg	1260
agttgatccg	1320
ccctcgggtt	1380
gggcctactc	1440
	1492

<210> 47  
 <211> 2406  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 8077668CB1

<400> 47	
atgtatgacc	caccaaaaagt
gtggaaatac	agtcagccct
taccgcgggt	caccgcgcg
cgcgctctgcg	tgggtggcca
ggctccgcgg	aggagccagt
gccagcctga	ccaaggatgc
atccctgctt	gctggagaat
gcccagcagg	atgccttggc
gagcctgtgc	tggttgggac
ggggttgtgc	tcttctctag
gactggctgg	ctggtgggaa
acacaggagt	gctcctgggc
	60
gatcctctcc	120
gggtgcgggtc	180
tgagggtagg	240
gagaggagaa	300
gatcaccagg	360
ccatgaggca	420
gttcatgtgg	480
caacatcttt	540
	600

atgttcctgg	tgtccttcgt	catcctgggt	gccctcgtca	cgggtgctgtc	tgacattggc	660
gtcggggagc	gcagcagcat	cggcagcggg	ggcgtctact	ccatgatctc	ctcgggtcctg	720
gggtgggcaga	cgggaggcac	catcgggctg	ctctatgtgt	ttggacagat	gtatatcacc	780
ggcttttgctg	aatccatctc	ggatttgctg	ggcctcggga	atatctgggc	tggtcgagga	840
atctcagttg	cgggtgcttct	ggccttgctg	ggcattaacc	tcgcaggtgt	caaattggata	900
atccgcctcc	agctgctggt	gctgttcctg	ctggccgtgt	ccacactgga	ctttgtgggtg	960
ggttctttca	cccacctgga	cccagaacat	ggtttcattg	gatattcacc	cgaactgcta	1020
cagaacaaca	cgctgcccca	ttacagcccg	ggggaatctt	ttttcactgt	ctttgggggt	1080
ttcttcccag	cggctacagg	agtcattggc	ggcttcaaca	tggggggcga	cctcagggag	1140
cctgccgcca	gcattccctc	gggctccctg	gcagctgttg	gcatctcgtg	gtttctgtac	1200
atcggctaca	ggagtcatgg	ccggcttcaa	catggggggc	gacctcaggg	agcctgccgc	1260
cagcattccc	ctgggctccc	tggcagctgt	tggcatctcg	taagttcccc	tcattggcaat	1320
cagtctgagc	tagagagtg	gagagaaaac	cccaggcaga	agagggattg	ccccctttgt	1380
tctgagagtg	cccgcactgt	gaccaatcta	tttcatattc	cgatccttct	tttaacactc	1440
gggtgggctg	aaaattcacg	tctgggtggg	tcagacactg	ccctttcatg	cgccctgcct	1500
gggaatgctg	ccgaggagga	gtgtcgggac	ccagatcaca	gatcgaatcc	tttctttcct	1560
ttctcgggcc	tcccggcccc	gtccctcccc	cacccccatt	ttgaacttgg	ggattttttt	1620
aaccaccgag	tttccaggat	cactggaggg	agcataaaga	tgaagagttt	tgcaatcaga	1680
cgaattgggg	ctgcattctc	gctctgcctc	tgattggctg	catgactctg	gacaatttac	1740
tgagttctct	gagcttccat	ttcctttctt	gtgaaatggg	gataatgaca	ctgtatgtgc	1800
ttcctagatt	gcggcgggga	gagaaggcac	acagtggag	ctcagcacc	agctcttctt	1860
cttccccag	tgagcctccc	ctccagctcc	cccaggccgt	caggcctctg	gggttttagt	1920
tcacactgtg	ccaccacac	aagcaggctg	ggttaacctg	gggtctatttc	tgacctttta	1980
gaatattggt	tctctctgta	ccacaggaag	aagcttctct	gcccttcatt	gggtaggtgt	2040
gaagaggcaa	agagatcaga	gaccagaaac	aatgacttga	gaagttaaaa	gggcctgtgt	2100
tattctaacc	acagtattta	gaaaatatta	atgtatctgt	ttcacaaaca	ctgaatttgt	2160
tcttagctta	ctgcagtga	accattgcta	tcagcctggc	gtacagcatc	aggggtatttt	2220
gtgcaataac	attgggaaga	aaatgaagac	aacagagata	gagttgagtg	ttgggaagga	2280
aataaaagaa	agaggcttct	gaatgaggac	cctgtcttgt	ttctggcccc	agctgggagg	2340
ttccaggagg	aaccttggc	ccatcctctg	attaccacga	tgggaatgag	gagaccacag	2400
gaacct						2406

<210> 48

<211> 3686

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 55120485CB1

<400> 48

atggaagaga	acagtaagaa	ggaccatcgg	gctttgctca	accaggagga	ggaggatgaa	60
ctggagggtg	ttggttaccg	ggaccacaat	gtacggaaag	ccttctgcct	tgctgcaccc	120
gtgctgacct	gtgggggccc	tctgtgtgtg	ttctactgga	gacccagtg	gagagtgtgg	180
gccaactgca	tcccatgccc	cttgcaagaa	gcagacactg	ttttgctgag	gacaacagac	240
gaatttcaaa	gatatatgag	gaagaaggta	ttctgcctct	acttatacac	actgaagttt	300
cctgtaagca	agaagtggga	agaatccctg	gtggctgacc	gccactctgt	cataaaccac	360
gccttaataa	agccagaatt	aaaactgcgg	tgccctggaag	tgcaaaaaat	caggatatgt	420
tggaaacgacc	tggagaagcg	gtttcagaaa	gttgggttgc	tagaagacag	caattcctgc	480
tctgacatcc	atcagacatt	tggattgggt	ctgaccagtg	aagagcaaga	ggtcagaaga	540
ttagtgtgtg	ggcccaacgc	cattgaggtt	gaaatccaac	ccatatggaa	gctgcttgtt	600
aaacagggtt	taaatccatt	ctatgtgttc	caagccttca	ccctaacttt	gtggctgtct	660
caaggttaca	tagaatactc	tgtggccatc	atcattttga	ctgttatctc	cattgtctta	720
agtgtgtatg	atttgcgaca	gcaatcagtt	aagctgcata	acctcgtgga	ggaccacaac	780
aaagtcagag	ttacaatcat	tgtaaaagac	aaagggttgg	aggagctgga	atcccgcttc	840
ttggttcccg	gagacattct	tattcttcca	ggaaaatttt	cattgccatg	tgatgctgtt	900
ttgattgatg	gaagctgcgt	ggtgaatgaa	ggcatgctta	caggagaaag	tatacctgtt	960
acaaagacac	cattgcccac	gatggagaac	actatgcctt	ggaaatgtca	cagtttggag	1020
gattatagga	aacacgtcct	tttctgtgga	acagaagtta	tccagggtcaa	gccctctggg	1080
caggggcctg	tacgagcagt	cgttttgcaa	acaggttaca	atacagccaa	aggggactta	1140
gtgagatcca	tctctgtacc	cggcctctg	aacttcaaac	tatacagcga	tgccctcaag	1200
ttcatcgtgt	tcttggcctg	ccttgggtgtc	atgggttttt	tctatgcctc	aggggtatat	1260

atgtaccatg	gagttcctcc	aaaagatact	gtgaccatgg	ccctgatcct	cctcaccgtg	1320
actgtccctc	cagtgtctgcc	agctgccctg	accataggca	acgtgtatgc	tcagaagaga	1380
ctgaagaaaa	agaaaatctt	ctgtatctcc	ccacagagaa	tcaacatgtg	tgggcaaata	1440
aacctcgtgt	gctttgacaa	aactggcact	ctcactgaag	atgggctgga	cctctggggg	1500
actgtcccta	ctgctgacaa	ctgcttccag	gaagcccaca	gctttgcctc	aggccaggct	1560
gtgccatgga	gcccactgtg	tgcggccatg	gccagctgcc	actctctgat	ccttctcaat	1620
gggaccatcc	agggagaccc	tctggacctc	aaaatgtttg	agggcactgc	ctggaaaatg	1680
gaagattgca	ttgtagactc	ctgcaaattt	gggacgtcag	tttcaaacat	cataaaacca	1740
ggaccaaaag	ccagtaagag	tccagtggaa	gccatcatca	ccttgtgcca	gtttccattt	1800
tcctcgagcc	tgcagaggat	gtccgtgac	gctcagctag	ctggggagaa	tcattttccat	1860
gtctacatga	aagtgcccc	agaaatggtg	gccaggttct	gcagatctga	aacagtgcc	1920
aagaatttcc	cacaggaact	gaggagttag	acggtgcaag	gcttccgtgt	cattgctctt	1980
gcccacaaaa	ccttaaagat	ggggaatctt	tcagaagtcg	agcacttagc	cagagaaaaa	2040
gtggagtcag	agttaacatt	tctgggactt	ctcatcatgg	agaatcgctt	gaaaaaagaa	2100
accaaactgg	tcttgaagga	actgagtgag	gcccgtatca	ggactgtgat	gattacaggt	2160
gataaccttc	aaacggccat	tactgttgca	aaagaattctg	aaatgatccc	tccaggcagc	2220
caagtgatca	ttgttgaggc	cgatgaacca	gaagaatttg	ttcctgcctc	tgtgacctgg	2280
cagctggtgg	agaaccaaga	gactggacct	gggaagaaag	aatctacat	gcatactgga	2340
aacagttcaa	ccctcgtgg	ggaaggagga	agctgttacc	atthtgcaat	gagtgggaaa	2400
tcataccaag	tgatatttca	gcattttcaac	agcttgcttc	caaaaattct	ggtgaatgga	2460
acagtttttg	caagaatgtc	tctctggcag	aaatcaagcc	ttattgaaga	atttcagaaa	2520
ttaatgcct	gcacagtaca	aatgaaagc	atctcagagt	taaccatgtc	tccaactgct	2580
ccagaaaaaa	tggaaagtaa	tagcaccttc	acaagttttg	agaacactac	agtctggttc	2640
ttgggaacaa	tcaactgtat	cactgtggct	cttgtgttct	ctaaaggaaa	accattttaga	2700
cagccaactt	atacaaaacta	tatatattgtc	cttgtgctga	taatacagct	tgggtgatgt	2760
ctattcttgc	tatttgctga	tataccagaa	ttatatagac	gtttggatct	gctctgcact	2820
cccgctcctgt	ggagggcctc	cattgtcatc	atgctcagct	tgaatttcat	tgtgtccctt	2880
gtggccgagg	aggtgtttat	tgaaaatcga	gccctgtgga	tgatgattaa	aagatgtttc	2940
ggctatcagt	caaaaagcca	gtatcggata	tggcagaggg	acttggcaaa	tgaccctagt	3000
tggcccccgc	taaaacaaac	ctcccactct	gacatgcccg	agtgtggcag	aggagtgtct	3060
tacagcaatc	cagtatttga	gagcaactct	gaaccaactct	gaaggacatg	tgatatccaa	3120
gttgatatta	cagagaaaca	atgacaaaag	ggaccagttt	cattgatttt	aagaaatgag	3180
actcttgtaa	catcagctgg	agttttgggg	ctacctatca	aatcatgggtg	acaaactaaa	3240
gtctgtttga	taaaatactg	tctacagaga	agaacattga	cccagaagag	tgctttcctc	3300
ttggcttctg	aaggtattag	aggcattcag	tgaaggcaag	aatgggccaa	gaatttctat	3360
gggccaagag	acaagcagcc	taaaagggaa	aaatatgtga	tctttcctct	taggacagga	3420
caatactttc	ctttttataa	aaatgagttg	aggggaagta	gtactgtgta	tactcacgga	3480
gttgtaattt	ttgaattttt	acttcacga	tggaaaaact	aatgcgtggg	aaacattaca	3540
ccaatcttta	ctgtctgggt	tctttttctc	tgtagtactt	ttcagatttg	aggctttttc	3600
ttctgtttct	ggccttacc	aacatattgt	taaggataag	cttagcctat	gctgcaataa	3660
aaaacaactc	tcaaaaaaa	aaaaaa				3686

<210> 49

<211> 2807

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3112883CB1

<400> 49

aagaattcgg	cacgaggcca	cagccgcggg	tgcgtggctg	tgcggcaagt	gagcgatttg	60
gtgaacagac	actcccttct	gtagccatgg	attgttacag	aacttacta	agcagttcct	120
ggattttacc	cactgtgatc	ctctgcttat	ttggtttttt	ctccatgatg	agaccctcag	180
aaccattcct	tatcccatat	ttatctggac	cagataaaaa	cctgaccagt	gcagagataa	240
caaatgagat	cttccccgtt	tggacatact	cctacctggg	gctgctgctg	cctgtgtttg	300
tcctcaccga	ttatgtccgc	tacaagccag	tcacatctt	gcaaggatc	agtttcatca	360
ttacctggct	gctgctgttg	tttggccaag	gagtgaagac	catgcagggt	gtagagttct	420
tctatgggat	ggtcaccgcc	gccgaggtgg	cctactacgc	ctacatatac	agcgtgggtc	480
gccccgagca	ctaccagaga	gtgagcggt	actgcaggag	cgtcacgctg	gccgcctaca	540
cagcagggtc	ggtgctgggt	caactcttgg	tatccctggc	gaacatgtcg	tacttttacc	600
tcaacgtcat	atccttggcc	tctgtctcgg	tggctttcct	tttctcactt	ttcctaccaa	660

tgcccaagaa	aagcatgttt	tttcatgcaa	aaccagcag	agaaataaag	aagtcacaa	720
gcgtgaatcc	agtatttagag	gaaactcacg	aagggtgaagc	accaggctgt	gaagagcaga	780
aaccacatc	agaaatactc	agcacttcag	ggaagctgaa	taagggccag	ctgaacagcc	840
tgaaccaag	caatgtgact	gtggacgttt	ttgtgcagtg	gttccaagat	ttgaaggagt	900
gctactcctc	aaaacgtctt	ttctactggt	ctctatggtg	ggctttcgcc	acagcagggt	960
ttaaccagggt	tttgaactat	gttcaaatcc	tgtgggatta	caaggcgcca	tccaagatt	1020
cttccatcta	taatggggcc	gtagaagcta	ttgcaacctt	tggaggggct	gtggctgcct	1080
ttgcagtggt	ttatgtgaaa	gtcaactggg	accttctggg	agagctggct	ctgggtggtct	1140
tctcagttgt	caatgccggg	tctttatttc	tcatgcatta	cacagccaat	atctgggcgt	1200
gctatgctgg	ctatttgata	ttcaagtcca	gctatatgct	tcttataacc	atagcagtat	1260
ttcagattgc	agttaatctg	aatgtggaac	gctatgcctt	ggtatttggg	atcaacacct	1320
ttattgcctt	ggtgattcag	accatcatga	ctgtgattgt	agtagatcag	agaggggtca	1380
acttgccagt	cagcattcag	tttttagttt	atgggagcta	ttttgcagta	attgctggaa	1440
ttttcctaata	gagaagcatg	tatattacct	actcaaccaa	atcccagaag	gatgtacaga	1500
gccctgctcc	aagtggagaat	ccagatgtgt	ctcaccagaa	ggaagagagt	aatatcatca	1560
tgtaacaaaa	actctaacct	catcgcaaca	aacgcaacag	tggctttcaa	agttatgcaa	1620
taataaggaa	agatttttag	atgggtggca	tatgttttgc	cataacttga	catgctttgc	1680
aaatctggat	tccaatggac	ctttcaaaac	cacaacaaaa	cctcagtttt	agatgagttc	1740
tctatgtgac	caattttact	ggatgcaatt	gacaggaccc	gtcatcataa	ttaaacaacc	1800
catattgggg	acccctctgt	actagtagca	gctggaaaat	tctggttttt	atcacttgta	1860
aagacatgca	gatggcgtgg	aaccaaacca	tgagaaaact	ccagccatcc	tggagttgat	1920
attcaccatt	tgtgaggaga	aatactaact	ggactgaccc	tattgctagg	cttaaatact	1980
tatttgatct	taccaaagaa	gtcaacacat	gggacctttg	tgtcacatga	accattttct	2040
ttctctctct	attaagtgta	tttctgttta	agttacagtt	ctctaagaga	attacaatgt	2100
ttgtcccatt	tctaagggct	tctcttcaac	tctaataaca	gcattattca	cgttatgatt	2160
tgatagtatt	attattttaat	ttttttatga	ttattttcca	ttttgtgtct	tgagttttgc	2220
tgttgaaagt	ctccctcaag	aatagcttca	gatectcttg	tgtatttgca	gaatacacaa	2280
ggtcatttcc	cagtggcctg	ggagaggcag	tgagccttct	ctccaccacc	atagacagggt	2340
gttaatgcat	ctatgggcca	ggtgcagtg	ctcactcctg	taatccgagc	aatttgggag	2400
gccaaagggtg	gagttttgct	tgaggctggt	ctcaaaactcc	tgacttcagg	tgatctgtct	2460
gcctcggcct	cccaaagagc	tgagattaca	ggcgtgagcc	accatgcctg	gccaaaacac	2520
ctgatttttta	aatggtagca	attgtgtctaa	ataacatgtg	tcaacaattt	tcaaagagtt	2580
gttaacattt	tttgatgaat	tctgtctgaaa	attatctgaa	catacctgtt	tttcgaaaact	2640
tcaagaaaact	ttcaataaag	gatgcattga	gccgtctgac	agtgagaata	acgatgctga	2700
gtcagatttt	attaggctgg	acttgtggtt	acaactttat	cccgcatttt	gggaagctaa	2760
gtgagaggta	attgaggttg	gggtttaaac	ccttggcatt	ttaacct		2807

<210> 50

<211> 2170

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4253888CB1

<400> 50

gtgaatgccc	cttccatccc	ccgcccctgg	tgtctcggtc	tggcttgagg	agacggggac	60
ccttctcaca	ccggcccgctg	ccgcccgcgc	gctccggaac	agatccagtc	cttctgtgga	120
acttctgaac	atctttttatt	agtggaaata	ttttctacac	aatgaagtca	acaacttaat	180
ttaaaccagt	gtttgtgcgg	ttctgattca	tctgctgtgg	ttcccgaagc	ttgagatcta	240
aggagtacag	ggtcttttgt	gatgacaata	tgactaatag	taaaggaaga	tctattaccg	300
ataaaacaag	tgggtggtcca	agtagtggag	gaggttttgt	agattggact	ttacgtttaa	360
acacaattca	atccgacaag	tttttaaatt	tactcttgag	tatggttcca	gtgatttacc	420
agaaaaacca	agaagacagg	cacaaaaaac	caaacggcat	ttggcaagat	ggattatcaa	480
ctgcagtaca	gactttttagt	aatagatctg	agcaacacat	ggagtatcac	agtttctcag	540
agcagtcctt	tcatggccaat	aatgggcacg	catcatcaag	ctgcagccaa	aagtatgatg	600
actatgccaa	ttataattac	tgtgatggaa	gggagacttc	agaaaccact	gccatgttac	660
aagatgaaga	tatatctagt	gatggtgatg	aagatgctat	tgtagaagtg	accccaaaat	720
taccaaagga	atccagtgge	atcatggcat	tgcataact	tgtgcccttt	ttgctagctg	780
gttttggaac	agtttccagct	ggcatggtag	tggatatagt	acagcactgg	gaggtgttca	840
gaaaagttac	agaagttttc	attttagtcc	ctgcacttct	tggctctcaa	gggaacttgg	900
aaatgacatt	ggcatccaga	ttatccactg	cagtaaatat	tgggaagatg	gattcaccca	960

ttgaaaagtg	gaacctaata	attggcaact	tggctttaaa	gcaggttcag	gcaacagtag	1020
tgggttttct	agcagctgtg	gcagcaatta	tattgggctg	gattccagaa	ggaaaatatt	1080
accttgatca	ttccataact	ctgtgctcta	gcagtgtggc	aactgccttc	attgcatctc	1140
ttctgcaggg	aataataatg	gttgggggta	tcgttgggtc	aaagaagact	ggtataaatc	1200
ctgataatgt	tgctacaccc	attgctgcta	gttttggcga	ccttataact	cttgccatat	1260
tggcttggat	aagtcagggc	ttatactcct	gtcttgagac	ctattactac	atttctccat	1320
tagttggtgt	atTTTTcttg	gctctaaccc	ctatttggat	tataatagct	gccaaacatc	1380
cagccacaag	aacagttctc	cactcaggct	gggagcctgt	cataacagct	atggttataa	1440
gtagcattgg	gggccttatt	ctggacacaa	ctgtatcaga	cccaaacttg	gttgggattg	1500
ttgtttacac	gccagttatt	aatgggtattg	gtggtaattt	ggtggccatt	caggctagca	1560
ggatttctac	ctacctccat	ttacatagca	ttccaggaga	attgcctgat	gaacccaaag	1620
gttgttacta	cccatttaga	actttctttg	gtccaggagt	aaataataag	tctgctcaag	1680
ttctactgct	tttagtgatt	cctggacatt	taattttcct	ctacactatt	catttgatga	1740
aaagtggta	tacttcttta	actataatct	tcatagtagt	gtatttattt	ggcgctgtgt	1800
tacaggtatt	taccttgctg	tggattgctg	actggatggg	ccatcacttc	tggaggaaag	1860
gaaaggaccc	ggatagtttc	tccatcccc	acctaacagc	attgggtgat	ctgctcgga	1920
cagctctgtt	agccttaagt	tttcattttc	tttggcttat	tggagatcga	gatggagatg	1980
ttggagacta	ataaattcta	caaactgctc	tcaagttacc	aaggaagaaa	atacacgaca	2040
accacttatg	gctctttttc	aaaactctta	aatcagtagt	ttgacttttg	ccagggtaat	2100
cttcagttgg	ccctgattca	attaaatggc	cttaattttt	ttttaaggaa	tttgtgtcaa	2160
aaaaaaaaa						2170

<210> 51

<211> 1722

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7479974CB1

<400> 51

atggatgctg	ttaaatactt	aaataaattg	aatttagata	acattgagtt	aacaaaatat	60
ttgtttttta	ctggttaaagg	tggcgtaggc	aaaacaacga	tatcaagttt	tattgcttta	120
aacttagcag	agaatggaaa	gaaagtagct	ttagtaagta	ctgatccagc	tagtaattta	180
caagatgtat	ttcaaatgga	attatcta	aaattaacta	aatatcaacc	tataccta	240
ctctctatag	ccaatttcga	cccgaattgt	gctgcagacg	attataaagc	acaatctata	300
gaaccttatg	agggtattct	accagaagat	gtgcttgctg	aaatgaaaga	acagtttaagt	360
ggttcatgta	cagttgaagt	agcagcattt	aatgaattta	caaatttttt	atccgataaa	420
actttagaac	aagaatttga	tttcattata	tttgatacag	ctccccacag	tcacactttg	480
agaatgcttg	aattacccttc	tgcatggaca	gattatttaa	atacaacgag	taatgacget	540
tcttgcttag	gtcaattatc	aggttttaat	gaaaatagag	ttaaatataa	ttcagcactt	600
gaaaaactac	gtaaccaaga	tgatacgacc	atgatgtag	ttgagagacc	tactcactct	660
tctatatatg	aaattcaaag	agcgcaacaa	gaattacaac	aactgtcaat	ttctaaattc	720
aaagtaatca	ttaacaacta	tatagaagaa	agtcacgggt	taatttcgag	tcagatgaaa	780
tcggaacaag	ataaaaaacat	taatcatttt	actgaatggg	taaataacaa	tcagtcttat	840
tacgttccat	ataaaaaagca	gaaagaagaa	ggtagagaaa	atttaactaa	tctattaaat	900
gatgataact	taattgaaaa	tgatgacttt	attgttgaag	atcatccgca	attcaataaa	960
ttaatcgatg	aaattgaaaa	tagtaaagtt	caatatttat	ttacaatggg	aaaagggtggc	1020
gttggtaaaa	cgacagtagc	aacgcaatta	gctacagcat	tatctaataa	aggatatcgt	1080
gttcttttag	caactactga	ccctactaaa	gaaattaatg	ttgaaactac	aagtaattta	1140
aatactgctt	atattgatga	agaacaagca	ttagaaaagt	ataaaaaaga	agtactagcc	1200
acagtgaatg	atgatacacc	acaagacgat	attgattata	ttatggaaga	tttaaaatca	1260
ccttgtagag	aagaaatagc	atTTTTcaaa	gccttttagt	acattatgga	gaatcaagac	1320
gacatggatt	acgtcattgt	agacacagct	cctacaggcc	ataccttgct	gttacttgat	1380
tctagtga	atcatcatag	agaattaaag	aaaaaatcaa	ctcaactac	cagtaagtgt	1440
gaaacattat	tacctaaaat	tcaaaaataa	aatTTaacac	agatgataat	cgtaacacta	1500
gcagaaaaaa	caccttat	agaatctaaa	cgtttagtag	aagattttaa	tagagcta	1560
ataggccata	attggtgggt	tgttaatcaa	tcgttagtta	cgctaaatca	acgtgatgac	1620
cttttttagta	acaaaaaaga	agatgaatca	ttttggataa	acaagattaa	aaatgaaagt	1680
tttgacaatt	atTTTgtcat	accttatggg	gggttatcat	aa		1722

<210> 52

<211> 1424  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7483850CB1

<400> 52  
 ctgatgtcta ttttaggaaga aacaaactgg tgtatatattg cttgttttcta catttttctag 60  
 gcactgcttt gttatctgtg gtggccttttc ccttggaaga tttctatctg ggcacttacc 120  
 atgcagttaa caacatacca aaaacagagc agcacaggct gttttatgta gcactgctga 180  
 ccatttgctt tggcattgga ggcgtaagag ccacgtctctg tccactgggt gcttttgccc 240  
 ttcaggagta tggatcacaa aaaacgatgt ctttttttaa ctgattcctt caggatatta 300  
 tctgcaaaact atgaattcca acctgaattt ggatggattt cttctgccga ttgcagtaat 360  
 gaatgccatc agcagcctac cattactaat tctggctcct tttctggagt atttcagcac 420  
 ctgcctgttt ccctctaaga gagttggatc atttctgtca acatgcatca ttgctggaaa 480  
 tctttttgct gcattgtctg tgatgatagc tggcttcttt gaaatacacc gaaaacattt 540  
 ccctgcagtg gagcagcccc tttcaggaaa agttctcact gtttccctcca tgccctgttt 600  
 ctacctgatt cttcagtatg ttttacttgg agtggcggaa acactggtaa acccagccct 660  
 ctctgtaata tcatacagat ttgttccaag caatgtcaga ggaacctcca tgaattttct 720  
 gacactgttc aatggatttg gctgtttcac aggggcactg ctggtgaagt tggatatctt 780  
 catctcagaa ggtaaaaata ggcaatgggt tccaaacaca ttaaacaag gcaatttaga 840  
 aggccttctt tcttctctgg catcattaac attgttgaac gtccctgggt tctgcagtgt 900  
 ttcacaaaaga tattgtaac taaatcattt taatgccag aacatccgtg gaagtaattt 960  
 tgaagaaaca cttctctcc acgaaaaatc tctgaaattt tatggcagta tacaggaatt 1020  
 ttcttcaagt attgatcttt gggagacagc cctatgaaac tgtgtttgag tctacctgtc 1080  
 ttatgagaac agtattcatt gttttcttct tcagttgagc attttatgtg ttttagtgta 1140  
 aaagccaata tatacagagg tgattttaca aatatcatct gtatacttta tttacaaaaga 1200  
 ctaattatga attttctagt aaactgtcat gcttttgcag taacatgtta tagcattaga 1260  
 atttaataaa ataatttcaa agtctcagtg attactctat gacaaaaaaa aaaggggggc 1320  
 gccgattagt gagcttcgtc gaccgggaaa taattccgga cggacctgca ggcgtaggcg 1380  
 atcctgagat tccatcaatg ggggcgtcga gcatgtttaa aggc 1424

<210> 53  
 <211> 3598  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5508353CB1

<400> 53  
 caatttgcaa cacagtagtg gtttctgctc ctaaccaacc cggacaaaag atcagacacc 60  
 cttcactggg ggggttgccc attaagtctt tggaaagat taaaagtctt ttccagagat 120  
 ggtctgtccg aagatcaagt tctccatcgc ttaacagtgg gaaagagcca tcttctggag 180  
 ttccaaacgc ctttgtgagc agactccctc tctttagtcg aatgaaaacca gcttcacctg 240  
 tggaggaaga ggtctcccag gtgtgtgaga gccccagtg ctccagtagc tcagcttgct 300  
 gcacagaaac agagaaacaa cacggtgatg caggcctcct gaatggcaag gcagagtccc 360  
 tccctggaca gccattggcc tgcaacctgt gttatgaggg cgagagccca gacgaagcgg 420  
 ccttagtgta tgccgccagg gcttaccat gcactttacg gtctcggaca ccagagcagg 480  
 tcatggtgga ctttgtgctt ttgggaccat taacatttca actcctacac atcctgcctt 540  
 ttgactcagt aagaaaaaga atgtctgttg tggtcggaca cctcttttcc aatcaagttg 600  
 tgggtgtatac gaaaggcgct gattctgtga tcatggagtt actgtcgggt gcttccccag 660  
 atggagcaag tctggagaaa caacagatga tagtaaggga gaaaaccag aagcacttgg 720  
 atgactatgc caaacaaggc cttcgtactt tatgtatagc aaagaaggctc atgagtgaca 780  
 ctgaatatgc agagtggctg aggaatcatt ttttagctga aaccagcatt gacaacaggg 840  
 aagaattact acttgaatct gccatgaggt tggagaacaa acttacatta cttgggtgcta 900  
 ctggcattga agaccgtctg caggaggagg tcctgaatc tatagaagct cttcacaag 960  
 cgggcatcaa gatctggatg ctgacagggg acaagcagga gacagctgtc aacatagctt 1020  
 atgcatgcaa actactggag ccagatgaca agctttttat cctcaatacc caaagtaag 1080  
 atgcctgtgg gatgctgatg agcacaattt tgaagaagct tcagaagaaa actcaagccc 1140

tgccagagca	agtgtcatta	agtgaagatt	tacttcagcc	tcctgtcccc	cgggactcag	1200
gggttacgagc	tggtactcatt	atcactggga	agaccttgga	gtttgcctcg	caagaaagtc	1260
tgcaaaagca	gttcctggaa	ctgacatctt	ggggccaagc	tggtgctctgc	tgccgagcca	1320
caccgctgca	gaaaaagtga	gtgggtgaaat	tggtccgcag	ccatctccag	gtgatgacct	1380
ttgctattgg	tgatgggtgcc	aatgatgtta	gcatagata	agtggcagac	attgggatag	1440
gggtctcagg	tcaagaaggc	atgcaggctg	tgatggccag	tgactttgcc	gtttctcagt	1500
tcaaacatct	cagcaagctc	cttcttgtcc	atggacactg	gtgttatata	cggctttcca	1560
acatgattct	ctatTTTTTt	tataagaatg	tggtctatgt	gaacctcctt	ttctgggtacc	1620
agttcttttg	tggatttttca	ggaacatcca	tgactgatta	ctgggttttg	atcttcttca	1680
acctcctctt	cacatctgcc	cctcctgtca	tttatgggtg	tttgagagaaa	gatgtgtctg	1740
cagagacctt	catgcaactg	cctgaacttt	acaaaagtgg	tcagaaatca	gaggcatact	1800
taccccatat	cttctggatc	accttattgg	atgcttttta	tcaaagcctg	gtctgtctct	1860
ttgtgcctta	ttttacctac	cagggtcag	atactgacat	ctttgcattt	ggaaaccccc	1920
tgaacacagc	cgctctgttc	atcggtctcc	tccatctggg	cattgaaagc	aagagtttga	1980
cttggattca	cttgcctggc	atcattggta	gcactctgtc	ttatttttta	tttgccatag	2040
tttttgaggc	catgtgtgta	acttgcaacc	caccatcaa	cccttactgg	attatgcagg	2100
agcacatgct	ggatccagta	ttctacttag	tttgtatcct	cacgacgtcc	attgctcttc	2160
tgcccagggt	tgtatacaga	gttcttcagg	gatccctgtt	tccatctcca	attctgagag	2220
ctaagcactt	tgacagacta	actccagagg	agaggactaa	agctctcaag	aagtggagag	2280
gggtctgaaa	gatgaatcaa	gtgacatcaa	agtatgctaa	ccaatcagct	ggcaagtcag	2340
gaagaagacc	catgcctggc	ccttctgtcg	tatttgcaat	gaagtcagca	agttcctgtg	2400
ctattgagca	aggaaactta	tctctgtgtg	aaactgcttt	agatcaaggc	tactctgaaa	2460
ctaaggcctt	tgagatggct	ggaccctcca	aaggtaaaga	aagctagata	ccctccttgg	2520
agttgcaagt	attctttcaa	ggttggaaga	gggattttga	agaggtatct	ctccaagcaa	2580
gaatgacttg	tttttccata	agggacatga	gcattttact	aggcttgga	gagctgacat	2640
gatgagcatt	attgtatggt	tgtatataca	tttgtgatag	agggttagag	tttgacctag	2700
agagagttaa	aggaagtga	atatttaatt	cagaaccaa	tgcttttgta	aaactttttg	2760
gattttgtaa	aagcattttc	attctcttag	aaattcaagt	attttcaagg	ggagtcatct	2820
gagatatatt	tattttacta	ggagatctta	tattctaggg	aaatgcttta	aatggtcagg	2880
ctccagctcg	aattttttta	agaaaaaagt	agttttta	acattgggta	ggactcagag	2940
gaaatacggg	aaaaacattg	tagatgggta	atttacagat	aaaatcccaa	gagcctttta	3000
aacaacaagg	tacctaaaat	aggggtataat	tatactgctt	aaaatacagt	tagtgcctat	3060
taatagcttt	ttatttccct	tggggagatg	cttttggtct	tctggctgaa	gatgtaggca	3120
tacctctcac	tcatttcaat	gttttctctga	gggtggagcct	tcattggaaa	ggggaaagag	3180
ggattctagg	gtttcatcag	ggaccaggaa	tgcatctctc	tgtaggttc	caatcaagag	3240
aagacctttt	atgagatctg	cctctgtata	gatgtcgta	gttaggaaac	tgaagccata	3300
ggtcaggaga	catcagctca	gcctgtggcc	atgggtgatt	tcctgtattt	taaaactgac	3360
agtagcctga	tccaacgtga	tatcaatcat	tttcaacacc	atccttcaga	ggacaatgtg	3420
aggtcattaa	gttttacaat	tacctggaga	ttttcagggtg	gtggggcctc	taaaattgga	3480
ttaacacgga	actatacagt	ggtcggagg	tcaaggggtc	tcgaaaattt	ggtacatctc	3540
tcggcgggag	tcacacagtc	cctcggactt	gtgcgggtgg	cacacattta	cagggggca	3598

<210> 54

<211> 1485

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 8543628CB1

<400> 54

gctattggat	tctccctata	tgagtcgtat	taatttctac	tacgaaatgc	tcgccgtgct	60
gacgctgatt	tgcgtgctgt	tggtcttgcc	gttgctcgctg	ttgctcagcc	gcctggaaag	120
gaggtccag	catggccagt	tcgggtcttg	agttgctctg	gggtgctgtg	ccgcaactgg	180
gcaagggcgc	tgcgcaaac	ctgtcaattt	cgtttttgag	catcggttc	agtacggttg	240
gcggcgtgct	ctacggcgtg	ctgcgcactt	tgaataacaa	ggcgatcaac	ggcgtgttg	300
gggtctacct	ggaactgttc	cgcgccattc	cgggtgctgg	gtgggtgtac	ctgctgtttt	360
ttggcgtgce	gatcttcttc	ggcctgagca	tcccgagttt	ctgggtgcgcg	gtgctggtgc	420
tgtcggtgtg	gggcgcagc	gaggtggcg	aagtgggtgcg	tggcgcgctg	cattcgctgc	480
cccggtggcca	acgcgaagca	ggcttgctga	ttggcctgtc	cgacctacag	ttgtacggct	540
acgtactgct	gccccaggcc	ctcaggcgca	tgacgcgcgc	gacctcaac	gtgtacacgc	600
gcatcatcaa	gaccagctcc	ctggcgggtgc	tgatcgccgt	gggtggatgtg	atcaaggctcg	660

```

gccagcaaat catcgagcgc acctacgagt cggatttgat ctacgggtgcg ctgttcctgt 720
ttttcttctt tatctgctac ccgttgctcg ccgcctccaa gctgctggaa cggcgctggg 780
cccaagcatg agcgcatgga tcgagttcca ggggttcaac aaattcttcg gcgagggcga 840
ggtgctcaag ggcacgcgacc tgagcgtgca gcgcggcgaa gtggtggtga tcttagggcc 900
cagcggctgt ggcaaaagca ccttggtgcg ctgcctcaat gggctggaag tcgcgcacag 960
cggcagcctg cgctttgccc gcaaagagct gctggataaa accaccgact ggcgccaggt 1020
acgccaggac attggcatgg tattccagag ctaccacctg tcccgcacat gagcgtgcct 1080
cgaaacatcc tgctcggccg ctgaaagtgg aaaagcgcgga ccgcgtgggt gcctgtgagc 1140
aggcggaaaa atgctgggag cgcctggccc ttgcccgaag cgcgacgcct tcccacgcc 1200
agctttccgg tggccagcaa caacgcacgc catcgtcctg cgttggtgcat gaacctcag 1260
tcatgctgtt gacgaagttc accggggcct cgcgaaatg gcaggaatgt tggagtgtcc 1320
agaattgccg gaggggtacc ctgtgtttcc cagaatggtt gccggcgcg cgacaggggtg 1380
atgagggcg tctcaaaaac ccggttttta ccgcacgcgc gactggagct tttccccc 1440
gaggaatta cgctgcctcg gggagaacaa cggagacaca cgcgc 1485

```

<210> 55

<211> 1470

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7482754CB1

<400> 55

```

gcctgttctg cagaggagga ggctgggctc agggcacttt tcacgcccc tccgttgggc 60
tggggacctt ggtaagccat ggaggacag accccaggaa gcaggggcct tccagaaaag 120
cctcaccctg ccacggctgc tgccactctg tctcgtatgg gcgctgtctt catcctcatg 180
aagtccgcgc tgggagctgg cctgctcaac ttcccctggg ctttctccaa agcgggcgga 240
gtggtccctg ccttccctgg ggagctggc tcgttggctt tctgatcag cgggctgggc 300
atcctgggct atgctgctgc tgtcagtggc caggccacct accagggtgt ggtcaggggg 360
ctgtgtggcc ctgccattgg gaagctgtgt gaggcctgct tctcctcaa cctgctcatg 420
atctccgtgg ccttcctcag ggtgatcggg gaccagctgg agaagctgtg tgactccctc 480
ctgtctggca ccccgccgc cccgcagccg tggtagcag accagcgctt caccctgccc 540
ctgctctccg tgetggctat cctgcccctg tctgccccgc gggagatcgc cttccagaaa 600
tacacaagcc cctcacagcg tcaactgtgt agcatcctag gcactctggc tgctgttac 660
ctggccctgg tcatcaccgt gcagtactac ctctggcccc agggcctcgt gcgtgagtcc 720
catccttcac tgagccctgc ctctggacc tctgtgttca gtgtcttccc caccatctgc 780
ttcgggttcc agtgtcacga agctgcccgc tccatctact gcagcatgcg caaacggagc 840
ctctcccact gggccctggt gtctgtgctg tcttgctgg cctgctgctt catctattca 900
ctgacggggg tttatggctt cctgactttt gggacagaag tttctgctga cgtcttgatg 960
tctacccag gcaatgatat ggtcatcatt gtggccccgg tcttttttgc tgtctccatc 1020
gtaactgtct accccatcgt gctcttctcg gggaggtcag tgatgcagga cttctggagg 1080
aggagctgct tggggggatg ggggcccagc gccctggccg acccctcagg gctgtgggtc 1140
cggatgccgc tgaccatcct gtgggtcacc gtgacgctcg ccattggcgt gtttatgcct 1200
gacctacgc agatcgtcag catcatcgga ggcacagt tcttcttcat cttcatcttc 1260
ccaggccctg gaggcaaaaca ccgtttctgc gaaggatggt ttcttaggaa accttcccaa 1320
acagcccaag cattatctga gaaggggaag ccagcaaggt ttgtgcctca tctgtgcaat 1380
gggtgtcgag cctataggac caagagtcaa gtgagaaacc tgcgagccac aggaagagag 1440
tctctgtcta cccagtatga gggcatctga 1470

```

<210> 56

<211> 3132

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3794818CB1

<400> 56

```

atggagtttg tgcgggcgct gtggctgggc ctggcgctgg cgctggggcc ggggtccgcg 60
gggggccacc ctacgccgtg cggcgctcct gcgcgcctcg ggggtccgt gcgcctgggc 120

```



```

gccctcctgc cccgcgcgcc tctcgcccgcc gccgcgcgcc gcgcgcgcct ggccccgggc 180
gccctggcgc cgcggctgcc gcacaacctg agcttgagc tgggtgctgc cgcgcgcgc 240
gcccgcgacc ccgcctcgtc gacccgcggc ctgtgccagg cgctggctgc tccgggcgtg 300
gcggccctgc tcgcctttcc cgaggctcgg cccgagctgc tgcagctgca cttcctggcg 360
gcggccaccg agacccccgt gctcagcctg ctgcgccggg aggcgcgcgc gccctcgga 420
gccccgaacc cattccacct gcagctgcac tgggccagcc cctggagac gctgctggat 480
gtgctgggtg cggtgctgca ggcgcacgcc tgggaagacg tcggcctggc cctgtgcgcg 540
actcaggacc ccggcgccct ggtggccctc tggacaagcc gggctggccg gccccacag 600
ctggtcctgg acctaagccg gcgggacacg ggagatgcag gactgcgggc acgcctggcc 660
ccgatggcgg cgccagtggg ggggtgaagca ccggtaccgc cggcggtcct cctcggctgt 720
gacatcgccc gtgcccgtcg ggtgctggag gccgtacctc ccggcccca ctggtgttg 780
gggacaccac tgcgcgccaa ggcctgccc accgcggggc tgccaccagg gctgctggcg 840
ctgggcgagg tggcacgacc ccgctggag gccgcatcc atgacattgt gcaactgggtg 900
gccccggcgc tgggcagtgc ggcacaggtg acccggaagc gagccctcct cccgcgcgcg 960
gtcaactgcg gggacctgca gccggccggg cccgagtccc cggggcgctt cttggcacgg 1020
ttcctggcca acacgtcctt ccagggccgc acgggccccg tgtgggtgac aggcagctcc 1080
caggtagaca tgtctcgcca ctttaaggtg tggagcctc gccgggaccc acggggcgcc 1140
ccggcctggg ccacggtggg cagctggcgg gacggccagc tggacttggg accgggaggt 1200
gcctctgcac ggcccccgcc cccacaggtt gccaggtct ggcccaagct gcgtgtggtg 1260
acgctgttgg aacacccatt tgtgtttgcc cgtgatccag acgaagacgg gcagtgcaca 1320
gcggggcagc tgtgcttggg ccctggcacc aacgactcgg ccacctgga cgcactgttc 1380
gccgcgctgg ccaacggctc agcgcgccgt gccctgcgca agtgctgcta cggctactgc 1440
attgacctgc tggagcggct ggcggaggac acgccttcg acttcgagct gtacctcgtg 1500
ggtgacggca agtacggcgc cctgcgggac ggcgctgga ccggcctggt cggggacctg 1560
ctggccggcc gggccacat ggcggtcacc agcttcagta tcaactccgc ccgctcacag 1620
gtgggtggact tcaccagccc cttcttctcc accagcctgg gcatcatggt gcgggcacgg 1680
gacacggcct caccatcgg tgcctttatg tggccctgc actggtccac gtggctgggc 1740
gtctttgcgg ccctgcacct caccgcgtc ttcctcaccg tgtacgagtg gcgtagcccc 1800
tacggcctca cgccacgtgg ccgcaaccgc agcacgtct tctcctactc ctcagccctc 1860
aacctgtgct acgccatcct cttcagacgc accgtgtcca gcaagacgcc caagtgcacc 1920
acgggcccgc tgtcatgaa cctctgggcc atcttctgcc tggctggtgt gtcagctac 1980
acggccaacc tggctgccgt catggtcggg gacaagacct tcgaggagct gtcggggatc 2040
cacgacccca agctgcacca ccgggcgacg ggttccgct tcggcaccgt gtgggagagc 2100
agcgcggagg cgtacatcaa gaagagcttc cccgacatgc acgcacacat gcggcgccac 2160
agcgcccca acacgccccg cggcgtcgcc atgctcacga gcgaccccc caagctcaac 2220
gccttcatca tggacaagtc gctcctggac tacgaggtct ccacgacgc cgactgcaaa 2280
ctgctgaccg tgggaaagcc cttcgccatt gagggctatg ggatcggact gccccagaa 2340
tcgcgctca cctccaacct gtccgagttc atcagccgct acaagtctc cggttcac 2400
gacctgtccc acgacaagtg gtacaagatg gtgccttgcg gcaagcgggt ctttgcggtt 2460
acagagaccc tgcagatgag catctaccac ttcgcgggcc tcttcgtgtt gctgtgcctg 2520
ggcctgggca gcgtctgct cagctcgctg ggcgagcacg ccttctccg cctggcgctg 2580
ccgcgcatec gcaaggggag caggctgcag tactggctgc acaccagcca gaaaatccac 2640
cgcgcctca acacggagcc accagagggg tcgaaggagg agacggcaga ggcggagccc 2700
agcggccccg aggtggagca gcagcagcag cagcaggacc agccaacggc tccggagggc 2760
tggaaacggg cgcgcggggc cgtggacaag gagcgccgcg tgcgcttct gctggagccc 2820
gccgtgggtg tggcaccga agcggaacgc gaggcggagg ctgcgcgcg agagggcccc 2880
gtctggctgt gctcctacgg ccgccccccc gccgcaaggc ccacgggggc cccccagccc 2940
ggggagctgc aggagctgga gcgcgcacat gaagtgcgc gtgagcggt cccagagccc 3000
ctggtgcggc gcggccagct cctggcacag ctcggggaca gcgcagctca ccggcctcgg 3060
cgcttgcttc aggcagagc gggccccgcg gagggccac cacactctgg ccgaccgggg 3120
agccaggaat ga 3132

```

<210> 57

<211> 1832

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4717525CB1

<400> 57

ggagctgacc ctgcgggggc ccgggggggg agggggagcc gcgaagcccc cactgagtgc 60

cgccgctgcc	gggggggggag	ggggagccgc	gaagccccc	ctgaggccgc	cgctgccggg	120
cctcccctcc	ccccggggcg	ggcgccatgc	gggggagccc	ggcgacgcgc	gagcgggcggc	180
agcgctgggg	tcgcctgttc	gaggagctgg	acagtaacaa	ggatggccgc	gtggacgtgc	240
acgagttgcy	ccaggggctg	gccaggctgg	gcgggggcaa	cccagacccc	ggcgcccaac	300
agggtatctc	ctctgagggg	gatgctgacc	cagatggcgg	gctcgacctg	gaggaatttt	360
cccgtatctc	gcaggagcgg	gaacagcgtc	tgctgctcat	gtttcacagt	cttgaccgga	420
accaggatgg	tcacattgat	gtctctgaga	tccaacagag	tttccgagct	ctgggcattt	480
ccatctcgct	ggagcaggct	gagaaaattt	tgcaacagcat	ggaccgagac	ggcacaatga	540
ccattgactg	gcaagaatgg	cgcgaccact	tcctgttgca	ttcgttgaa	aatgtggagg	600
acgtgctgta	tttctggaag	cattccacgg	tcctggacat	tggcgagtg	ctgacagtg	660
cggacgagtt	ctcaaagcaa	gagaagctga	cgggcatgtg	gtggaagcag	ctggtggccg	720
gcgcagtggc	aggtgccgtg	tcacggacag	gcacggcccc	tctggaccgc	ctcaaggctc	780
tcacgaggt	ccatgcctca	aagaccaacc	ggctgaacat	ccttgggggg	cttcgaagca	840
tggtccttga	gggaggcatc	cgctccctgt	ggcgcgga	tggtattaat	gtactcaaga	900
ttgccccga	gtcagctatc	aagttcatgg	cctatgaaca	gatcaagagg	gccatcctgg	960
ggcagcagga	gacactgcat	gtgcaggagc	gcttcgtggc	tggtccctg	gctggtgcca	1020
cagcccaaac	catcattttac	cctatggaga	ctctgaagaa	ctggtggctt	cagcagtaca	1080
gccacgactc	ggcagaccca	ggcatcctcg	tgctcctggc	ctgcggtacc	atatccagca	1140
cctgcygcca	gatagccagt	taccgctgg	ccctggtccg	gaccccgatg	caggcacaag	1200
cctccatcga	gggtggcccc	cagctgtcca	tgctgggtct	gctacgtcac	atcctgtccc	1260
aggaggggcat	gcggggcctc	taccggggga	tcgcccccaa	cttcatgaag	gttattccag	1320
ctgtgagcat	ctcctatgtg	gtctacgaga	acatgaagca	ggccttaggg	gtcacgtcca	1380
ggctcgagta	cagtggctcg	atctcagatc	actgcaacct	ctgccttcca	ggttcaagcg	1440
attctcctgc	ctcagcctcc	agagtagctg	gtattacagg	atttcatcat	gttgcccagg	1500
ctcatctcgg	actcgtgggt	tcaaggaatt	cggcagcctt	cagccttcca	acgtgctggg	1560
attacaggaa	gccggtcgct	tgcacatgag	cagccttatg	gagtggaacca	tgtggttaagg	1620
aactcagcca	atagccatgt	aactgagctt	tgaagaggat	cttgetgtcc	tggccaacat	1680
ctcactgcga	ttctatcagt	tgaatttcct	ggatagtcca	agctttgttg	gatccctgca	1740
ccggtacagc	tgggttccag	taactgaatc	ctgagttcta	tgatgtttta	cttcaagcat	1800
gggtcagggg	acagcgggtt	atgggttcct	tg			1832

<210> 58

<211> 1902

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5091793CB1

<400> 58

cccaccttca	gtcctggctc	tgacaagccc	tccagcttca	cgccaccg	gatgggagaa	60
agcaggtgtc	gcgagagttg	ggcgcaagac	gccttgtagg	gagtgtaact	atggccggcc	120
tgcggaacga	aagtgaacag	gagccgctct	taggcgacac	acctggaagc	agagaatggg	180
acattttaga	gactgaagag	cattataaga	gccgatggag	atctattagg	attttatata	240
ttactatggt	tctcagcagt	gtaggggttt	ctgtagtcat	gatgtccata	tggccatata	300
tccaaaagat	tgatccgaca	gctgatacaa	gttttttggg	ctgggttatt	gcttcatata	360
gtcttggcca	aatggttagct	tcacctatat	ttgggtttatg	gtctaattat	agaccaagaa	420
aagagcctct	tattgtctcc	atcttgattt	ccgtggcagc	caactgcctc	tatgcatata	480
tccacatccc	agcttctcat	aataaatact	acatgctggg	tgctcgtgga	ttgttgggaa	540
ttggagcagg	aaatgtagca	gttggttagat	catatactgc	tggtgctact	tcccttcagg	600
aaagaacaag	ttccatggca	aacataagca	tgtgtcaagc	attaggtttt	attctagggtc	660
cagtttttca	gacttgtttt	acattccttg	gagaaaaagg	tgtgacatgg	gatgtgatta	720
aactgcagat	aaacatgtat	acaacaccag	ttttacttag	cgccttccctg	ggaattttta	780
atattattct	gatccttgcc	atactaagag	aacatcgtgt	ggatgactca	ggaagacagt	840
gtaaaagtat	taattttgaa	gaagcaagta	cagatgaagc	tcagggtccc	caaggaaata	900
ttgaccagggt	tgetgttggt	gccatcaatg	ttctgttttt	tgtgactcta	tttatctttg	960
ccctttttga	aaccatcatt	actccattaa	caatggatat	gtatgcctgg	actcaagaac	1020
aagctgtgtt	atataatggc	ataatacttg	ctgctccttg	ggttgaagcc	gttgttattt	1080
tcttaggagt	taagttgctt	tccaaaaaga	ttggcgagcg	tgctattcta	ctgggaggac	1140
tcacgttgt	atgggttggc	ttctttatct	tgttaccttg	gggaaatcaa	tttcccaaaa	1200
tacagtggga	agatttgcac	aataattcaa	tccctaatac	cacatttggg	gaaattatta	1260
ttggtctttg	gaagtcctca	atggaagatg	acaatgaaag	accaactggt	tgctcgattg	1320

aacaagcctg	gtgcctctac	accccggtga	ttcatctggc	ccagttcctt	acatcagctg	1380
tgctaataagg	attaggctat	ccagtcctgca	atcttatgtc	ctatactcta	tattcaaaaa	1440
ttctaggacc	aaaacctcag	ggtgtataca	tgggctggtt	aacagcatct	ggaagtggag	1500
cccggattct	tgggcctatg	ttcatcagcc	aagtgtatgc	tcactgggga	ccacgatggg	1560
caatcagcct	ggtgtgtgga	ataatagtgc	tcaccatcac	cctcctggga	gtgggtttaca	1620
aaagactcat	tgctctttct	gtaagatatg	ggaggattca	ggaataaact	agctaagact	1680
gtgatggaaa	ctacttgctg	tgtggcactt	cctggctctaa	agctctgcta	gacaattgctg	1740
gttgggtgca	gtgtctcatg	cctgtaatcc	cagcactttg	gaaggccaag	gcaggaggat	1800
tgcttgaggc	caggaattca	agaccagcct	gggaaacaaa	gtgagaccct	gtctttacaa	1860
aaaataaaaa	caaaaattta	aaactgtaaa	aaaaaaaaaa	aa		1902

<210> 59

<211> 2820

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5945527CB1

<400> 59

gcgcccagct	ctgtagcctc	ctccgtcgac	tcagccttag	gtaccggtea	ggcaaaatgc	60
ggtcctccct	ggctccggga	gtctcggttct	tccgggcctt	ctccagggac	agctgggtcc	120
gaggcctcat	cctgctgctg	accttcctaa	tttacgcctg	ctatcacatg	tccaggaagc	180
ctatcagtat	cgtcaagagc	cgtctgcacc	agaactgctc	ggagcagatc	aaacctatca	240
atgatactca	cagtcctcaat	gacaccatgt	ggtgcagctg	ggccccatct	gacaaggaca	300
actataagga	gttactaggg	ggcgtggaca	acgccttctc	catcgccctat	gccatcggca	360
tgttcatcag	tgggggtttt	ggggagcggc	ttccgctccg	ttactacctc	tcagctggaa	420
tgctgctcag	tggccttttc	acctcgctct	ttggcctggg	atatttctgg	aacatccacg	480
agctctggta	ctttgtggtc	atccaggctc	gtaatggact	cgtccagacc	acaggctggc	540
cctctgtggg	gacctgtggt	ggcaactggt	tgggaaaggg	gaagcggggg	ttcatcatgg	600
gcatctggaa	ttccacacac	tctgtgggca	acatcctggg	ctccctgatc	gccggcatct	660
gggtgaacgg	gcagtggggc	ctgtcgttca	tctgtccctg	catcattact	gccgtcatgg	720
gcgtcatcac	cttcctcttc	ctcatcgaa	acccagaaga	tgtggactgc	gccccctctc	780
agcaccacgg	tgagccagct	gagaaccagg	acaacctga	ggacctggg	aacagtcctc	840
gctctatcag	ggagagcggc	cttgagactg	tggccaaatg	ctccaagggg	ccatgcgaag	900
agcctgctgc	catcagcttc	tttggggcgc	tccggatccc	aggcgtggtc	gagttctctc	960
tggtgtctgt	gtttgccaag	ctgggtcagtt	acaccttcc	ctactggctg	cccctctaca	1020
tcgccaatgt	ggctcacttt	agtgcgaagg	aggctgggga	cctgtctaca	ctcttcgatg	1080
ttgggtggat	cataggcggc	atcgtggcag	ggctcgtctc	tgactacacc	aatggcaggg	1140
ccaccacttg	ctgtgtcatg	ctcatcttgg	ctgcccccat	gatgttcctg	tacaactaca	1200
ttggccagga	cgggattgcc	agctccatag	tgatgctgat	catctgtggg	ggcctgggtca	1260
atggcccata	cgcgctcatc	accactgctg	tctctgctga	tctggggact	cacaagagcc	1320
tgaagggcaa	cgccaaagcc	ctgtccacgg	tcacggccat	cattgacggc	accggctcca	1380
taggtgcggc	tctggggcct	ctgctggctg	ggctcatctc	ccccacgggc	tggaaacaatg	1440
tcttctacat	gctcatctct	gccgacgtcc	tagctgctt	gtccctttgc	cggttagtat	1500
acaaagagat	cttggcctgg	aagggtgtccc	tgagcagagg	cagcgggtat	aaagaaatat	1560
gaggcccca	ttggaacagc	agcatggagg	gtccagttg	ggtccccaac	gtgctcccca	1620
tgggcaagac	aatggaaact	tccacaagca	ggggaggcaa	accctcttta	ttgaacatta	1680
gccagcccag	cccagacccc	agggctgcct	aaggacacag	agattctcca	tgggaagggg	1740
actgccaagc	atgaggaaat	agaagattca	ggggcctgag	ctctggaagc	tgcaagcaaa	1800
agggatggga	ctaggggctga	gttgtgtctc	catcttgata	aggaaaggat	atgctcacac	1860
tcttgcttgt	tcagattcca	agacagaagg	cttcacaagg	ccaacgcctg	gaaaatgggc	1920
atctctccct	cccatgttaa	gctttaacct	ctgtaatctg	cctgtatcta	taggtgggca	1980
tctcactcca	tcaaaggagc	ccagcctctc	tttgtccctc	taccatgcaa	cagtcttctc	2040
tgctgacttc	cccaagctgg	gccctcttct	actctccatt	taggcctggt	gataactcca	2100
ttaccacccc	atcactgctg	ttcctccagg	gccagcactc	gggcgaggca	ggggagctgc	2160
cttcggtaca	taatttgaag	gggcactccc	tcttggggcac	atgccggccc	tgagtgcctc	2220
ccttgccctca	ctctgatccc	ggccccataa	tgtcctcagt	ggaaggtgac	tggggggccgg	2280
tgctgtgggg	agagtagaaa	gaggggttgg	catgactaaa	aataaccagta	tgtgtattaa	2340
gtatttttag	aatgaaatgc	caaggagtgc	ctactgtatg	ccagctctgc	tctaggaatg	2400
gagtagacag	tggacacaag	aaggacttac	ggcctgagca	caggtgccaa	cggtgacaag	2460
actggcaaga	cgtgagggca	tgaatgggtc	attcaggcag	ctgctgcagg	tgtgggtcac	2520

tggtgccatc	tgtgtctccc	ttttccactt	ttctatgtcc	tccttccacc	ccaagtcccg	2580
gatcactcgc	tgttttctgg	ctagctcttg	gcattctccat	ctgagccctaa	agttgcccac	2640
tggcaccaat	agattctgtt	tgacctgctg	tgcccatgct	catctttgtc	ggggcttggc	2700
ttggcagcga	gtgaggctgg	catggggtgc	tgaagtgggg	tctcagttag	gggtgacagc	2760
gtgctggcag	tcctcacagc	cctcgtctgc	tctcggcgcc	tcctctgcct	gggctcccac	2820

<210> 60  
 <211> 3920  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6941124CB1

<400> 60	
ccggccgctga	gtgcgcgagt
gccgtgcggt	gggtgcctga
ccgtctcccg	tgggggaccc
ccggtgggcy	agagggtcgc
gcccgggcgc	ggcttctcgg
tcttgaattt	cttaccaggt
aagtgggcag	ggcggagctg
tggcgggccg	ccgcctcggg
gcgcgcggt	gggagccccc
gccctgcgcg	gccgcagcgg
cccgcacagg	cccaggtttg
ggagtgtgca	ccggcccaac
agcgcacct	agcagatggg
gctatggggc	ccctcccca
gttgcatccc	tccgaagagc
cacaggaggt	ccagggtctg
actctcccac	catgtcgcag
cagatgacac	cttcgaagcc
actcctccat	cgtgtctggc
ggcccacaac	actgagaatg
aaagtacctg	aagttaaaga
cttcctggcg	caggccagca
gcggaaccatg	cttcgcgcgt
ggacctgctc	atggccatgc
cctctgttca	gataccatcc
gtcgtggctc	tgcattcatc
ccgcctggtt	cgccacagag
ggtgctggcc	ccacccaaga
tgccaccatg	ttctcggata
ggaattcaag	gaggccttgg
agtggcgccg	agaacgaagg
gccccaaaag	tgcaaggact
caggttcccc	ttgtaccctt
gggcaaatac	atcaccacca
tttcgggtct	ctcaatgacg
cgggcagagc	atcgggggccc
gctgaccacc	gcgcccctgg
tgacctggac	ttcaactcct
gctttatgcc	tttttcaacc
gatcatcgcc	ctcttcatct
aatcttctgg	aagtactact
atcccttgtc	agcctgtcag
caatgccagc	ttcctcgcca
gaccgcgctg	ctcagcctcc
ccaattcaag	aagagccctc
cctgcccatc	gcgggtcctg
gatgagcaag	ttccgctaca
gtgcgcgagt	gtgcatgggc
gccgcgggaa	ggccgcccgg
gcggatgccg	tctgcctgcg
gccgcgggga	ggagggcgcg
cttggtttca	ttagggctct
gatgcagcgg	gtcaggggcy
agcgcctcca	gcgcagagct
ttaccttccg	cccgggcctc
agttttggga	gtcccgggac
gctctgacac	gggagggcca
tacctgcacg	tgggcccggc
cctcgtcgcc	cacggcctgg
gaattttctga	gattaaggct
ctgcttccct	ttctccctag
gccaggtcgg	ggggcgggga
ctggtgacct	ttctgcttcc
tcgaggattc	aagctactac
tgaggccttc	gacactgcca
gtttcttgtc	aatgtcaacc
gtgctcctgc	acacctcccc
gcgcaccgcy	acctagatgg
acctccctgg	ataacgtgct
aatgagccca	actgcaacct
cccattgogg	gtaaaagtcca
acagggggtgc	ggtaccagca
aagcggcacg	tgtgcatcag
gaggttcggg	tcgtcatcct
atggagggtgg	cgcgcacgtt
ctggagaccc	gcacagagga
accatgggtga	gccacgggtcc
cctgccccaca	gacacccaga
atccgggagg	acatcgcacg
attgggaaaa	acaaggctgt
tgccctcctgc	ccaccatcgc
gacgtgcaga	agaccatagc
gggcagccat	tgggtgattct
cgtgtcatct	gtgatgacta
tgggaatagtt	tcttctctgc
ttcaagagggt	cgacggagga
gccgtcaagg	gcacgggttaa
tatcacacaa	aaaggacttc
gccagcctcc	acactgccct
gccacacact	caggccaggc
tggctgggct	acaccctcta
gagatcctgt	ccgactgcgc
catggcttcc	gggaaatcga
gcgatggcgc	agatccagtc

gctgtccctg	agggccgta	gcggtgcca	gggcctcggc	ttcctgctgt	ccatgctctt	2880
cttcatcgag	cagaacttgg	tggccgcctt	ggtgaatgca	ccggagaaca	ggctggtgaa	2940
gggcactgcc	taccactggg	acctcctgct	cctcgccatc	atcaacacag	ggctgtctct	3000
gtttgggctg	ccttggatcc	atgccgccta	ccccactcc	ccgctgcacg	tgcgagccct	3060
ggccttagtg	gaggagcgtg	tggagaacgg	acacatctat	gacacgattg	tgaacgtgaa	3120
ggagacgcgg	ctgacctcgc	tgggcgccag	cgtcctgggtg	ggcctgtccc	tgttgetgct	3180
gccgggtccc	cttcagtggg	tccccaagcc	cgtgctctat	ggcctcttcc	tctacatcgc	3240
gctcacctcc	ctcgatggca	accagctcgt	ccagcgcgtg	gccctgctgc	tcaaggagca	3300
gactgcgtac	ccccgcacac	actacatccg	gagggtgccc	cagaggaaga	tccactactt	3360
caaggccctg	caggtgcttc	agctgctgct	gctgtgtgcc	ttcggcatga	gctccctgcc	3420
ctacatgaag	atgatctttc	ccctcatcat	gatcgccatg	atccccatcc	gctatatect	3480
gctgccccga	atcattgaag	ccaagtactt	ggtgtcatg	gacgctgagc	acaggccttg	3540
actggcagac	cctgcccacg	ccccattcgc	cagccctcca	cgtcctccca	ggctggctct	3600
ggagctgtga	ggggagggtg	aggtgtgtgg	gtgactgctc	tgtgctgcgc	cttctcatgg	3660
ctgactcagg	cctggggcat	ctgggcattg	taggggtgca	gtggtatgtg	cccacccctc	3720
tcccattatc	cttttagcttt	aggccaagag	cgttgctcag	ggcagcttct	gccagggtg	3780
ggtgggactg	agcaggatgg	attttctttt	gataaaagag	tcgatgcctg	aaagagaaac	3840
catttccttg	attgtgtaag	gaacttgctg	gacgcacatt	agagaataaa	gctcctgttt	3900
ctaggctcct	aaaaaaaaaa					3920

<210> 61  
 <211> 1333  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6972530CB1

<400> 61	
tggggcaacc	tgctaacagc
ttccttccgc	ggcagggcgt
attgcagcca	aggtacttct
ttgagcatcc	agacgccgag
ggctaaatgt	cgattttcct
ggttgggcag	gcctgatcgt
gtgttgctgg	gcgaacaaca
accacactca	ccgtcggcta
gtggcggcct	gggtcatgct
accagcagtg	ttatcgatgc
aaggctggcc	ataccgtggt
ttgttgctgc	aagacgaaac
gaagaaaatc	cgatgccggg
gcactgttgc	tgcgtgccgg
gatgacctga	cgctggccac
gtcgccact	tcaacgaaag
gagtgcacgt	ccagcatggc
tcggtggtca	tcaatgaatt
ttgccgaag	cctttgaggc
aacgccaccc	tcattggcta
agtgccacag	aggtcaaagg
atctcccatg	ggatggctta
ctcgagccga	acc
cgctgttgtg	gctttgctgt
ttccttcgaa	gggcatctgt
atccacgcgc	cgcagcgggt
gcgatgctgt	cgtcaggatt
actgctacgc	atgtacgcct
tgccttgggt	gtccacctga
cctcaccgct	gccgccacgt
tggcgacctc	gcgccgcaga
cgctgctgca	cgctgctgca
ggcatgaagg	gcaaagggtga
gatcggtgct	gagggcgcgt
ctcgaacgac	aacctgatcg
aaacgtgacg	tgatctgcga
caaagcagcg	ttcatccgcg
gtgccccggt	gccgagcgtg
gtgaaccaac	tgagtcccgt
gctcctacgc	gctcctacgc
cctgtgcgca	cctctcagga
ccaccagta	cctgatgaaa
cgcagatgaa	ggagcgccac
agccgctgat	caacccgccc
cttccactcg	ccttaaggaa
gcccacaaag	ccactaccga

<210> 62  
 <211> 8487  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6991750CB1

<400> 62

atgcacat	taataattta	cattcaattt	cagattgcag	aggagggatc	aaccattttct	60
tgtgtggtt	agagaaccag	aggggctctg	gattatgtgc	atgtttttta	caccattttca	120
cagattgaaa	ctgatggcat	taattacctt	gttgatgact	ttgctaattgc	cagtgggaact	180
attacattcc	ttccttggca	gagatcagag	gttctgaata	tatatgtttct	tgatgatgat	240
attcctgaac	ttaatgagta	tttccgtgtg	acattgggtt	ctgcaattcc	tgagatggg	300
aagctaggct	caactcctac	cagtgggtgca	agcatagatc	ctgaaaagga	aacgactgat	360
atcaccatca	aagctagtga	tcatccatat	ggcttgctgc	agttctccac	agggtgcct	420
cctcagccta	aggacgcaat	gaccctgcct	gcaagcagcg	ttccacatat	cactgtggag	480
gaggaagatg	gagaaatcag	gttattgggtc	atccgtgcac	agggacttct	gggaagggtg	540
actgcggaat	ttagaacagt	gtccttgaca	gcattcagtc	ctgaggatta	ccagaatgtt	600
gctggcacat	tagaatttca	accaggagaa	agatataaat	acattttcat	aaacatcact	660
gataattcta	ttcctgaact	ggaaaaatct	tttaaagttg	agttgttaaa	cttgggaagga	720
ggagccagtc	taggagtggc	ttcccaaatt	ctagtgacaa	ttgcagcctc	tgaccacgct	780
cattggcgat	ttgaatttag	ccctgagtca	ctctttgtca	gtggaactga	accagaagat	840
gggtatagca	ctgttacatt	aaatgttata	agacatcatg	gaactctgtc	tccagtgaact	900
ttgcattgga	acatagactc	tgatcctgat	ggtgatctcg	ccttcacctc	tggaacatc	960
acatttgaga	ttgggcagac	gagcgccaat	atcactgtgg	agatattgcc	tgacgaagac	1020
ccagaactgg	ataaggcatt	ctctgtgtca	gtcctcagtg	tttccagtgg	ttctttggga	1080
gctcatatta	atgccacggt	aacagttttg	gctagtgtatg	atccatatgg	gatatttcatt	1140
ttttctgaga	aaaaacagac	tgtaaagtt	gaggaagcaa	cccagaacat	cacactatca	1200
ataataagggt	tgaaggccct	catgggaaaa	gtccttgtct	catatgcaac	actagatgat	1260
atggaaaaac	caccttattt	tcacctaatt	ttagcgagag	caactcaagg	aagagactat	1320
ataccagctt	ctggatttgc	tctttttgga	gctaatacaga	gtgaggcaac	aatagctatt	1380
tcaattttgg	atgatgatga	gccagaaagg	tccgaatctg	tctttatcga	actactcaac	1440
tctacttttag	tagcgaaagt	acagagtcgt	tcaattccaa	attctccacg	tcttgggctt	1500
aaggtagaaa	ctattgcgca	actaattatc	attgccaatg	atgatgcatt	tggaactctt	1560
cagctctcag	caccaattgt	ccgagtggca	gaaaatcatg	ttggaccctt	tatcaatgtg	1620
actagaacag	gaggagcatt	tgcatgtgtc	tctgtgaagt	ttaaagctgt	gccaataact	1680
gcaatagctg	gtgaagatta	tagtatagct	tcatcagatg	tggtcttgct	agaaggggaa	1740
accagtaaaag	ccgtgccaat	atatgtcatt	aatgatattc	atcctgaact	ggaagaatct	1800
tttcttgtgc	aactgatgaa	tgaacaaca	ggaggagcca	gactaggggc	tttaacagag	1860
gcagtcatta	ttattgaggc	ctctgatgac	ccctatggat	tatttgggtt	tcagattact	1920
aaacttattg	tagaggaacc	tgagtttaac	tcagtgaagg	taaacctgcc	aataattcga	1980
aattctggga	cactcggcaa	tgttactggt	cagtgggttg	ccaccattaa	tggaacgctt	2040
gctactggcg	acctgcgagt	tgtctcaggt	aatgtgacct	ttggccctgg	ggaaaccatt	2100
caaaccttgt	tgtagagggt	cctggctgac	gacgttccgg	agattgaaga	ggttatccaa	2160
gtgcaactaa	ctgatgcctc	tggtggaggt	actattgggt	tagatcgaat	tgcaaatatt	2220
attattcctg	ccaatgatga	tccttatggt	acagtagcct	ttgctcagat	ggtttatcgt	2280
gttcaagagc	ctctggaaag	aagttcctgt	gctaataata	ctgtcaggcg	aagcggaggg	2340
cactttgggtc	ggctgttgtt	gttctacagt	acttccgaca	ttgatgtagt	ggctctggca	2400
atggaggaag	gtcaagattt	actgtcctac	tatgaatctc	caattcaagg	ggtgcctgac	2460
ccactttgga	gaacttggat	gaatgtctct	gccgtggggg	agccctgtta	tacctgtgcc	2520
actttgtgcc	ttaaggaaaca	agcttgctca	gcgttttcat	ttttcagtgc	ttctgagggt	2580
ccccagtggt	tctggatgac	atcatggatc	agcccagctg	tcaacaattc	agacttctgg	2640
acctacagga	aaaacatgac	cagggtagca	tctcttttta	gtggtcaggc	tggtggctggg	2700
agtgactatg	agcctgtgac	aaggcaatgg	gccataatgc	aggaagggtga	tgaattcgca	2760
aatctcacag	tgtctattct	tcttgatgat	ttcccagaga	tggtatgagag	ttttctaatt	2820
tctctccttg	aagttcacct	catgaacatt	tcagccaggt	tgaanaatca	gccaaccata	2880
ggacagccaa	atatttctac	agttgtcata	gcactaaatg	gtgatgcctt	tggaagtgtt	2940
gtgatctaca	atattagtcc	caatacttcc	gaagatggct	tatttgttga	agttcaggag	3000
cagcccaaaa	ccttgggtgga	gctgatgata	cacaggacag	ggggcagctt	aggtcaagtg	3060
gcagtcgaat	ggcgtgttgt	tggtggaaca	gctactgaag	gttttagattt	tataggtgct	3120
ggagagattc	tgacctttgc	tgaagggtgaa	accaaaaaga	cagtcatttt	aaccatcttg	3180
gatgactctg	aaccagagga	tgacgaaagt	atcatagtta	gtttgggtga	cactgaagggt	3240
ggaagtagaa	ttttgccaaag	ctccgacact	gttagagtga	acattttggc	caatgacaat	3300
gtggcaggaa	ttgttagctt	tcagacagct	tccagatctg	tcataggtca	tgagggagaa	3360
attttacaat	tccatgtgat	aagaactttc	cctggctcag	gaaatgttac	tggttaactgg	3420
aaaattattg	ggcaaaatca	agaactcaat	tttgctaact	ttagcggaca	acttttcttt	3480
cctgaggggt	cgttgaatac	aacattgttt	gtgcatttgt	tggtatgacaa	cattcctgag	3540
gagaaagag	tataccaagt	cattctgtat	gatgtcagga	cacaaggagt	tccaccagcc	3600
ggaatcgccc	tgcttgatgc	tcaaggatat	gcagctgtcc	tcacagtaga	agccagtgat	3660
gaaccacatg	gagtttttaa	ttttgctctt	tcatcaagat	ttgtgttact	acaagaggct	3720

aacataacaa	ttcagctttt	catcaacaga	gaatttggat	ctctaggagc	tatcaatgtc	3780
acataatacca	cggttcctgg	aatgctgagt	ctgaagaacc	aaacagtagg	aaacctagca	3840
gagccagaag	ttgattttgt	ccctatcatt	ggctttctga	ttttagaaga	aggggaaaca	3900
cgagcagcca	tcaacattac	cattcttgag	gatgatgtac	cagagctaga	agaatatttc	3960
ctggtgaatt	taacttacgt	tggacttacc	atggctgctt	caacttcatt	tcctccaga	4020
ctagattcag	aaggtttgac	tgcacaagtt	attattgatg	ccaatgatgg	ggccccaggt	4080
gtaattgaat	ggcaacaaag	caggtttgaa	gtaaatgaaa	cccatggaag	tttaacattg	4140
gtagccacaga	ggagcagaga	acctcttggc	catgtttcct	tatttgtgta	tgctcagaat	4200
ttggaagcac	aagtggggct	ggattatata	ttcaccccaa	tgattcttca	ttttgctgat	4260
ggagaaaggt	ataaaaaatgt	caatatcatg	attcttgatg	atgacattcc	agaaggagat	4320
gaaaaatttc	agctgatttt	aacaaatcct	ctcctgggac	tagagctagg	gaaaaataca	4380
atagccttaa	ttattgtcct	tgctaattgat	gacggccctg	gagttctatc	atttaacaac	4440
agtgagcact	tttccctaag	agagccaaca	gctctctacg	tcaggagag	tggtgcagta	4500
ttgtacattg	ttcggaacc	tgcacaagga	ttgtttggaa	cagtgcagct	tcagttcatt	4560
gtgacagaag	tgaattcctc	aaatgaatct	aaagatctga	ctccttccaa	aggctatatt	4620
gttttagaag	aaggtgttcg	attcaaggcc	ctacaaatat	ctgccatatt	agacacggaa	4680
ccagaaatgg	atgagtattt	tggtttgcac	ttgtttaatc	caactggagg	tgctagacta	4740
ggggtgcatg	ttcaaaccct	gataacagtt	ttgcaaaaacc	agggcccttt	ggggctatc	4800
agtatctctg	cagttgaaaa	tagagccacc	tccatagaca	tcgaagaagc	caataggacc	4860
gtgtatttaa	atgtatctcg	aactaatggc	attgatttgg	ctgatttaaa	tatagaaaaat	4920
ccgaaaactt	gtgagccctt	taatatgtgt	ttttctccct	actttgtgat	tactcatgaa	4980
gaaagaaatg	aagaaaagcc	ttctcttaac	agtgtgttta	cattcacatc	tggtatttaa	5040
ttattcctgg	tacaaacaat	cattattctg	gaaagttctc	aagtaagata	ttttacttca	5100
gacagccaag	attattttaat	cattgcaagt	caaagagatg	attccgaatt	aactcaggtc	5160
ttcaggtgga	atggaggaag	cttcgtgttg	catcaaaaac	tcctgtccg	aggtgtgctg	5220
accgtgcctt	tgttcaacaa	gggaggctct	gtgtctcttag	ccatttccca	ggctaagtcc	5280
aggctaaact	cccttttatt	cagatggtct	ggcagtggtg	ttattaactt	tcaagaggtg	5340
cctgtcagtg	ggacaacaga	agttgaggct	ttgtcttcag	ccaatgatat	ttaccttaata	5400
tttgcacaaa	atgtctttct	aggagatcag	aattcaattg	atattttcat	ctgggagatg	5460
ggcagctctt	ccttcaggta	ttttcagctc	gtagattttg	ctgctgttaa	cagaatccac	5520
tccttcacac	cagcctcagg	aatagcccac	atacttctta	ttggccaaga	tatgtctgct	5580
ctttactgct	ggaattcggg	gcgtaatcaa	ttctcttttg	ttctgggaagt	accttctgct	5640
tatgatgtgg	cttctgttac	agtaaagtcc	cttaattcaa	gcaagaattt	aatagctcta	5700
gtgggagctc	attcacatat	atatgagcta	gcctacattt	ccagccattc	tgactttatt	5760
cctagtccag	gtgaactgat	atttgaacct	ggtgagagag	aagctacaat	agcagtaaat	5820
atccttgatg	atacagttcc	agaaaaagaa	gaactcctca	aagttcaact	taaaaatccc	5880
aaaggaggag	cagagatttg	cattaatgat	tctgtaacaa	taaccattct	gtctaattgat	5940
gatgcctatg	gaattgtttg	atttgtctcag	aattcattat	ataagcaagt	ggaagaaatg	6000
gagcaagata	gcctagtaac	cttgaacgtt	gaacgcttaa	aaggaacata	tggccgtata	6060
accatagcat	gggaagctga	tggaaagtatt	agtgatatat	ttcctacctc	aggagtgggtg	6120
gaaaaacgaa	tgctagcaaa	gatttttatt	actgaaggcc	aggtactgtc	aacaatcact	6180
ctaactattc	ttgctgataa	tataccagag	ttatcagagg	ttgtgattgt	aacctcacc	6240
cgtatcacca	cagaaggggt	tgaggactca	tacaaagggtg	ctactattga	tcaggacaga	6300
agcaagtctg	ttataacaac	tttgcccaat	gactcacctt	ttggcttgggt	gggctggcgt	6360
gctgcgtctg	tcttcattag	agtagcagag	cctaaagaaa	acaccaccac	tcttcagtta	6420
caaatagtct	gagataaagg	actacttggg	gatattgcca	ttcacttgag	agctcaacc	6480
aatttcttac	tgcatgtcga	taatcaagct	actgagaatg	aagattatgt	attgcaagaa	6540
acaataataa	taatgaaaga	aaacataaaa	gaagctcatg	ccgaagtttc	cattttgccg	6600
gatgaccttc	ctgaattgga	ggaaggattt	attgtcacta	tcaactgagg	gaacctgggtg	6660
aactctgact	tctctacagg	acagccaagt	gtgcggaggc	ccggaatgga	aatagctgag	6720
ataatgatag	aagaaaatga	cgatcccaga	ggaattttta	tgtttcatgt	tactagaggc	6780
gctgggggaag	ttattactgc	ctatcaggtg	cctccaccct	tgaacgttct	tcaagttcct	6840
gtagtccggc	tggtctggaag	ctttggggca	gtaaatgttt	attggaaagc	atcaccagac	6900
agtgtcggcc	tggaagactt	taaaccatct	catgggattc	ttgaatttgc	agataaacag	6960
gttactgcaa	tgatagaaat	caccataatt	gatgatgctg	aatttgaatt	gacagagacg	7020
ttcaatattt	ccttgatcag	tgttgtctgga	ggtggcagac	ttggtgatga	tggtgtggtg	7080
actgttggtt	ttccacaaaa	tgattctcca	tttggagtat	ttggatttga	agaaaagact	7140
gtaactgattg	atgaatccct	ttcatccgat	gacctgatt	catatgtgac	attgacggtt	7200
gtccgggtccc	caggaggaaa	aggaaccgtc	cgacttgagt	ggaccataga	tgagaaggct	7260
aaacataacc	ttagtccctt	gaatgggacc	cttcattttg	atgagactga	gtcccagaag	7320
accattgtgt	tgcacacact	tcaagacaca	gtgttggagg	aggacaggcg	tttcaccatt	7380
cagctgatat	caattgatga	ggtagaaata	tctccagtaa	aaggtagtgc	atcaataatt	7440
attcgggggtg	ataagcgagc	atcaggagaa	gttgggtag	ctccgtcatc	taggcacatc	7500

ctcattgggg	aaccctcagc	aaaatataat	ggtaccgcta	ttatcagcct	tgttcgaggc	7560
ccagggat	tgggggaggt	cacagtgttc	tggaggatat	tccctccttc	cgtgggggaa	7620
tttgcgtgaaa	catcaggaaa	actgacaatg	cgagacgaac	agtctgcagt	cattgtagta	7680
atacagggtt	tgaacgatga	cattccccgag	gaaaaaagct	tctatgagtt	tcagctcact	7740
gcagtcaagt	agggaggagt	tctgagtga	tccagcagca	ctgccaacat	cacgggtggtg	7800
gccagcgact	ctccctatgg	ccgatttgcc	ttttcacatg	agcaacttcg	agtgtcagaa	7860
gcacagagg	ttaacatcac	aatcatccgt	tccagtggag	attttggcca	tgtgcgactc	7920
tggtagaaga	cgatgagcgg	gacagcggaa	gcaggcttgg	attttgttcc	tgcagcagg	7980
gagctcctct	ttgaagcagg	ggagatgagg	aaaagtctgc	atgttgaaat	ccttgatgat	8040
gactatcctg	aaggcccaga	ggaattttct	ctaacaatta	caaagggtgga	actccaggga	8100
aggtaaagga	gaaaggcaat	taggaaaaag	aaagcaaaga	gcagagagaa	agagagaaag	8160
ttctattgtg	tagagattaa	tactgagctc	taaagcagtt	ctttccctac	tctttctcag	8220
atttgtgtga	cattagaaca	ggacatacaa	gggaaattat	gacaatgact	gagttttttt	8280
ttttaatttt	gttaaatata	acttccctct	gaaaattgaa	agtacgtgtt	tctatttgtgt	8340
ggacacactc	tgggtagcca	ggaagttgaa	tatgataacg	gcttctaatt	ggtgaagaca	8400
cagcctaaca	gctctcta	tcagaggtca	gctggtgtgg	ctgaatttgc	ttcttttgat	8460
ttccatactt	gtgaaagggt	gaaggga				8487

<210> 63

<211> 3264

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 71726948CB1

<400> 63

cgctggaagc	agctggagge	tccagttctc	atctgctcag	gtgtccccgg	cgccttggcg	60
aactcgccca	ctccagttcc	tcacgtgggtg	agcactcagg	gcagcgggtc	gattttccga	120
ggtcccatac	ctgggtttga	ggggcgcggc	tcgcagcggc	gggtgcagg	gcgactgcca	180
gccctcaccc	cgcctcgagg	tgcttcgga	ggccgacacc	tggaggacgc	ctccagttcc	240
cgcgggaagc	cacgcctgcg	cgccagggat	ccgggataag	aagtgcgcgc	cgggctcccg	300
ctgcgcgcgc	cggggccacc	agtttgcgcg	cagggtctcag	gcgaccgtgc	ggccatggac	360
acgccacggg	gcacggcac	cttcgtgggtg	tgggactacg	tgggtgttcgc	gggcatgctg	420
gtactctcgg	cgcgcacgg	catctactac	gccttcgctg	ggggcgccca	gcagacctcc	480
aaggacttcc	tgatggcg	ccgcagaatg	accgcagttc	ccgtggcgct	gtccctcacc	540
gctagcttca	tgtcagccgt	cactgtcctg	ggcaccctct	ccgaggtcta	ccgttttggg	600
gccattttta	gcactcttgc	cttcacctac	ttctttgtgg	tggatcatcag	cgcggaggtc	660
ttctctcccg	tgttctacaa	actgggaatt	accagcacct	acgagtattt	agaacttcga	720
tttaacaaat	gtgttcgtct	ctgtggaaca	gtcctcttca	ttgttcaaac	aattctgtat	780
actggaattg	ttatttatgc	ccctgccctg	gctttgaatc	aagtcacagg	atttgatctg	840
tggggcgcg	tagtggcaac	gggggtgggtc	tgcacattct	actgcacact	gggtgggtctt	900
aaagcagtta	tctggaacaga	tgtttttcaa	gttgggatca	tgggtggctgg	atttgcattcc	960
gtgattatac	aggctgtgggt	gatgcaagggt	ggaatcagca	ctatttttaa	tgatgcctat	1020
gatggtggaa	gattaaattt	ctggaatttt	aatcctaacc	ctttgcaaag	acacaccttc	1080
tggacaatta	ttataggagg	gaccttcaca	tggaccagca	tctacgggtg	caaccaatcc	1140
cagggtgcaga	gatataattt	ttgtaaaagc	agattccagg	caaaactgtc	tctctacatc	1200
aatcttgttg	gactctgggc	aatcctcaca	tgctcagttg	tttgtgggct	cgccttatat	1260
tccaggtagc	atgactgtga	tccttggaca	gccaagaaag	tgtctgcacc	agaccagctc	1320
atgccttatt	tggtagctga	cattctgcaa	gattatccag	gacttctctg	actttttgtg	1380
gcctgtgctt	acagtgggac	attaaagcaca	gtgtcctcca	gtattaatgc	cttagcagca	1440
gtaactgtgg	aagatcta	caaaccttac	ttcagatcgc	tctcagaaag	gtctctgtct	1500
tggatttccc	aaggaaatgag	tgtggtgtat	ggagccctgt	gtattggaat	ggctgcgctg	1560
gcgtcaacta	tgggagcttt	gttcagggca	gcactcagcg	tatttgggtat	ggttgggtgga	1620
ccacttatgg	gcctgttcgc	tttgggcatt	ttgggtccct	ttgccaactc	aattggagca	1680
cttcttggtg	tgtatggctg	atttgccatt	tctctatggg	ttggaattgg	agctcaaata	1740
tatectccac	ttcctgagag	aacattgcca	ttgcaccttg	atatccaagg	ctgtaacagc	1800
acctacaatg	agacaaattt	gatgacaacc	acagaaatgc	catttactac	tagtgttttt	1860
caaataatac	atgttcaaag	gactccactg	atggataact	ggtattcttt	atcatatctg	1920
tacttcagca	ctgttggaa	tttggtaaca	ttattagttg	ggataactgt	cagtttatca	1980
acaggaggaa	gaaaacagaa	cttagacccc	agatatatac	taaccaaaga	ggacttttta	2040
tccaattttg	atatttttaa	gaaaaagaag	catgttttga	gctataaatc	acatccagtg	2100